

## REMARKS/ARGUMENTS

Prior to the present amendment, Claims 58-70 were pending in this application. With this amendment, Claims 64-68 have been canceled without prejudice and Claims 58-63 have been amended. Claims 58-62 and 69-70 are pending after entry of the instant amendment. The specification has been amended to correct formal errors as discussed below. The amendments to the specification and claims are fully supported by the specification and claims as originally filed and do not constitute new matter. Applicants expressly reserve the right to pursue any canceled matter in subsequent continuation, divisional or continuation-in-part applications.

### **I. Specification**

The disclosure was objected to because, according to the rejection, the correct priority date has not been submitted, while the specification of related application (U.S. Application Serial No. 09/999,829) has been amended to recite the correct priority information. Applicants would like to draw the Examiner's attention to the Preliminary Amendment submitted on August 21, 2002, presenting the correct priority information for the present application. The priority information recited in the Preliminary Amendment is identical to that of U.S. Application Serial No. 09/999,829. Thus, the correct priority information has been submitted and this objection should be withdrawn.

The disclosure was also objected to for containing errors in regard to SEQ ID NOs: 505 and 506. In particular, the Examiner stated that "on page 101, lines 16 and 17 and 31-32, the PRO213-1 of SEQ ID NO: 506 is stated as having 295 amino acids, but in the sequence listing, SEQ ID NO: 506 has 273 amino acids." The Examiner further stated that "on page 309, lines 10-12, the start and stop codons recited for SEQ ID NO: 505 is incorrect." Applicants have amended the specification to recite that "the PRO213-1 of SEQ ID NO: 506 has ~~295~~ 273 amino acids" and have changed the start codon of SEQ ID NO: 505 from "336-338" to "398-401" and the stop codon of SEQ ID NO: 505 from "1221-1223" to "1220-1222". The amendments are supported by the specification as originally filed, and do not constitute new matter. Support for the amendments can be found in Figure 212 and Figure 213.

As requested by the PTO, Applicants have amended the specification to correct the ATCC address on page 372, line 34. Further, the paragraph beginning at page 374, line 32, has

been amended to comply with the provisions of the Budapest Treaty. In addition, the specification has been amended to remove all embedded hyperlinks and/or other forms of browser-executable codes.

## **II. Double Patenting**

The Examiner alleges that "there are a series of applications in which SEQ ID NO:506 is present but do not claim the polypeptide." The Examiner has requested that Applicants point out to the Examiner all double patenting issues.

To our best knowledge, Applicants have not filed any applications having claims directed to a polypeptide of a sequence identical to SEQ ID NO: 506. Applicants believe that the Examiner reached his conclusion of the existence of possible conflicting claims based on the disclosure of the **publications** of other U.S. applications filed by Applicants, which do not reflect the changes made in preliminary amendments in those applications.

## **III. Priority**

The PTO asserts that Applicants are entitled to the priority of the filing date of the present application, October 15, 2001 only since the current application is not enabling for the nucleic acid of SEQ ID NO: 505. In particular, The PTO alleges that the exact same sequence has been given two different names (PRO213-1 and PRO1330) and is duplicated in the sequence listing. The PTO further alleges that confusing information exists regarding the gene amplification data for the various molecules. For example, the value given for PRO213 in Table 3 of the provisional application Serial No. 60/131,445 is identical to that given for PRO213-1 of the present application, therefore it appears that the values of Table 9 of the present application for PRO213-1 are actually the values for PRO213. In addition, it appears that the gene amplification data of PRO1330 (PRO213-1) of the present application is same as that of PRO213 of Table 3 of 60/131,445, and therefore the gene amplification data of PRO1330 (PRO213-1) is not present in the instant application. *See* page 4 of the instant Office Action.

The duplicated sequence (SEQ ID NO: 508) has been deleted in the revised sequence listing. Therefore, the present application is in sequence compliance now. In addition, Applicants respectfully submit that PRO213-1 was incorrectly designated as PRO213 in U.S.

Provisional Application Serial No. 60/131,445. Therefore, the  $\Delta C_t$  value for PRO213 in 60/131,445 is actually the  $\Delta C_t$  value for PRO213-1. PRO1330 is identical to PRO213-1, and therefore, the  $\Delta C_t$  value for PRO1330 is the same as that of PRO213-1 in the present application and also the same as that of PRO213 of 60/131,445. In addition, PRO213 and PRO213-1 are actually the same molecule, both being the same polypeptide encoded by the full-length coding sequence of DNA30943. Due to sequencing errors, the amino acid sequence of PRO213 differs from the sequence of PRO213-1 in a few positions. PRO213-1 is the correct sequence. Since the PRO213 and PRO213-1 polypeptides are in fact the same molecule, it is not surprising that they display the same biological function. Since the gene amplification data for PRO213-1 is disclosed in both the present application and provisional application U.S. Provisional Application Serial No. 60/131,445, claims directed to the PRO213-1 polypeptides are clearly supported by the disclosure of U.S. Provisional Application Serial No. 60/131,445.

Applicants rely on the gene amplification assay for patentable utility which was first disclosed in U.S. Provisional Application Serial No. 60/131,445, filed April 28, 1999, priority to which has been claimed in this application. Accordingly, the present application is entitled to at least an effective filing date of April 28, 1999.

#### **IV. Claim Rejections Under 35 U.S.C. §§ 101 and 112, First Paragraph (Enablement)**

Claims 58-70 stand rejected under 35 U.S.C. §101 allegedly "because the claimed invention is not supported by either a substantial and specific asserted utility or a well established utility." (Page 4 of the instant Office Action). Claims 58-70 are further rejected under 35 U.S.C. §112, first paragraph allegedly because one skilled in the art would not know how to use the claimed invention "since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility." (Page 9 of the instant Office Action). The Examiner specifically notes that "the nucleic acids have utility as cancer marker.... However, the protein does not have any specific and substantial utility, or a well-established utility" and therefore concludes that no asserted utility is specific for PRO213-1 protein. The Examiner also asserts that the data showing the amplification of the nucleic acids encoding PRO213-1 is not indicative of a use of the encoded polypeptide as a diagnostic or therapeutic agent. Further, the Examiner alleges that since the data are not corrected for

aneuploidy, and because it does not necessarily follow that an increase in gene copy number results in increased gene expression, the data do not support the implicit assertion that PRO213-1 can be used as a cancer diagnostic. The Examiner further quotes exemplary references like Pennica *et al* and Gygi *et al*. to show that "it does not necessarily follow that an increase in gene copy numbers results in increased gene expression and increased protein expression, such that antibodies would be useful diagnostically or as target for cancer drug development." For the reasons outlined below, Applicants respectfully disagree.

Applicants submit that the cancellation of Claims 64-68 renders the rejection of this claim moot. With respect to Claims 58-63 and 69-70, Applicants submit, as discussed below, that not only has the PTO not established a *prima facie* case for lack of utility, but that the claimed polypeptides possess a specific and substantial asserted utility.

#### **Utility – Legal Standard**

According to the Utility Examination Guidelines ("Utility Guidelines"), 66 Fed. Reg. 1092 (2001) an invention complies with the utility requirement of 35 U.S.C. §101, if it has at least one asserted "specific, substantial, and credible utility" or a "well-established utility."

Under the Utility Guidelines, a utility is "specific" when it is particular to the subject matter claimed. For example, it is generally not enough to state that a nucleic acid is useful as a diagnostic without also identifying the conditions that is to be diagnosed.

The requirement of "substantial utility" defines a "real world" use, and derives from the Supreme Court's holding in *Brenner v. Manson*, 383 U.S. 519, 534 (1966) stating that "[t]he basic quid pro quo contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with substantial utility." In explaining the "substantial utility" standard, M.P.E.P. §2107.01 cautions, however, that Office personnel must be careful not to interpret the phrase "immediate benefit to the public" or similar formulations used in certain court decisions to mean that products or services based on the claimed invention must be "currently available" to the public in order to satisfy the utility requirement. **"Rather, any reasonable use that an applicant has identified for the invention that can be viewed as providing a public benefit should be accepted as sufficient, at least with regard to defining a "substantial" utility."** (M.P.E.P. §2107.01, emphasis added.) Indeed,

the Guidelines for Examination of Applications for Compliance With the Utility Requirement, set forth in M.P.E.P. §2107 II (B) (1) gives the following instruction to patent examiners: "If the applicant has asserted that the claimed invention is useful for any particular practical purpose . . . and the assertion would be considered credible by a person of ordinary skill in the art, do not impose a rejection based on lack of utility."

Finally, the Utility Guidelines restate the Patent Office's long established position that any asserted utility has to be "credible." "Credibility is assessed from the perspective of one of ordinary skill in the art in view of the disclosure and any other evidence of record . . . that is probative of the applicant's assertions." (M.P.E.P. §2107 II (B) (1) (ii)) Such a standard is presumptively satisfied unless the logic underlying the assertion is seriously flawed, or if the facts upon which the assertion is based are inconsistent with the logic underlying the assertion (Revised Interim Utility Guidelines Training Materials, 1999).

The PTO also sets forth the evidentiary standard as to utility rejections. In general, an Applicant's assertion of utility creates a presumption of utility that will be sufficient to satisfy the utility requirement of 35 U.S.C. §101, "unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope." *In re Langer*, 503 F.2d 1380,1391, 183 USPQ 288, 297 (CCPA 1974). See, also *In re Jolles*, 628 F.2d 1322, 206 USPQ 885 (CCPA 1980); *In re Irons*, 340 F.2d 974, 144 USPQ 351 (1965); *In re Sichert*, 566 F.2d 1154, 1159, 196 USPQ 209, 212-13 (CCPA 1977).

Compliance with 35 U.S.C. §101 is a question of fact. *Raytheon v. Roper*, 724 F.2d 951, 956, 220 USPQ 592, 596 (Fed. Cir. 1983) cert. denied, 469 US 835 (1984). The evidentiary standard to be used throughout *ex parte* examination in setting forth a rejection is a preponderance of the totality of the evidence under consideration. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). Thus, to overcome the presumption of truth that an assertion of utility by the applicant enjoys, the Examiner must establish that it is more likely than not that one of ordinary skill in the art would doubt the truth of the statement of utility. Only after the Examiner made a proper *prima facie* showing of lack of utility, shifts the burden of rebuttal to the applicant. The issue will then be decided on the totality of evidence.

### **Proper Application of the Legal Standard**

As discussed below under the section on "priority", Applicants rely on the gene amplification data for patentable utility for the claimed polypeptides.

Gene amplification is an essential mechanism for oncogene activation. The gene amplification assay is well-described in Example 114 of the present application, the inventors isolated genomic DNA from a variety of primary cancers and cancer cell lines that are listed in Table 9, including primary lung cancers of the type and stage indicated in Table 8 (page 546). As a negative control, DNA was isolated from the cells of ten normal healthy individuals, which was pooled and used as a control. Gene amplification was monitored using real-time quantitative TaqMan™ PCR. The gene amplification results are set forth in Table 9. Further, Example 114 explains that the results of TaqMan™ PCR are reported in  $\Delta C_t$  units, wherein one unit corresponds to one PCR cycle or approximately a 2-fold amplification relative to control, two units correspond to 4-fold amplification, 3 units to 8-fold amplification etc. PRO213-1 showed approximately 1.03-5.55  $\Delta C_t$  units which corresponds to  $2^{1.03}$ - $2^{5.55}$  fold (more than 2 fold) amplification in human lung and colon tumors, which is significant and thus the PRO213-1 polypeptide or a portion thereof, such as a polypeptide comprising amino acid residues 35-273, has utility as a diagnostic marker of lung or colon cancer.

It is well known that gene amplification occurs in most solid tumors, and generally is associated with poor prognosis.

In support, Applicants submit a Declaration by Dr. Audrey Goddard with this response and particularly draw the Examiner's attention to page 3 of the declaration which clearly states that:

It is further my considered scientific opinion that an at least **2-fold increase** in gene copy number in a tumor tissue sample relative to a normal (*i.e.*, non-tumor) sample is significant and useful in that the detected increase in gene copy number in the tumor sample relative to the normal sample serves as a basis for using relative gene copy number as quantitated by the TaqMan PCR technique as a diagnostic marker for the presence or absence of tumor in a tissue sample of unknown pathology. Accordingly, a gene identified as being amplified at least 2-fold by the quantitative TaqMan PCR assay in a tumor sample relative to a normal sample is **useful as a marker for the diagnosis of**

**cancer**, for monitoring cancer development and/or for measuring the efficacy of cancer therapy. (Emphasis added).

The attached Declaration by Audrey Goddard clearly establishes that the TaqMan real-time PCR method described in Example 114 has gained wide recognition for its versatility, sensitivity and accuracy, and is in extensive use for the study of gene amplification. The facts disclosed in the Declaration also confirm that based upon the gene amplification results, one of ordinary skill would find it credible that PRO213-1 or a portion thereof is a diagnostic marker of human lung or colon cancer.

Secondly, regarding the Examiner's point that "none of [the] asserted utilities is specific for the disclosed PRO213-1 proteins or antibodies," Applicants submit, as discussed below, that the Examiner has not established a *prima facie* case for lack of utility for the claimed polypeptides.

**A prima facie case of lack of utility has not been established**

The Examiner bases the assertion, that increases in gene copy number do not reliably correlate with increased gene expression or polypeptide expression, on exemplary literature reports like Pennica *et al.* and Gygi *et al.* and hence concludes that the PRO213-1 polypeptides and their antibodies lack utility.

According to the Examiner, Pennica *et al.* teaches that "An analysis of *WISP*-1 gene amplification and expression in human colon tumors **showed a correlation between DNA amplification and over-expression**, . . . . In contrast, *WISP*-2 DNA was amplified in colon tumors, but its mRNA expression was significantly reduced in the majority of tumors compared with expression in normal colonic mucosa from the same patient." (Emphasis added). Firstly, Applicants draw attention to Pennica's showing that "a correlation between DNA amplification and over-expression exists for the *WISP*-1 gene" in 84% of the tumors examined. While Pennica discloses a lack of correlation for the *WISP*-2 gene, Pennica teaches nothing regarding such a lack of correlation in genes in general. That is, Pennica's teachings are specific for the *WISP* family of genes, and are not directed to genes in general. The Utility Guidelines requires that for a *prima facie* showing of lack of utility, the Examiner has to provides evidence that it is **more likely than not** that a lack of correlation between protein expression and gene amplification

exists, in general. Accordingly, Applicants respectfully submit that Pennica teaches nothing of the correlation between gene amplification and polypeptide over-expression in general.

Further, the Examiner cites the Gygi *et al.* reference to establish that "even if gene amplification correlates with increased transcription, it does not always follow that protein levels are also amplified." The Examiner adds that "Gygi *et al.* studied 150 proteins... and found no strong correlation between proteins and transcript levels." Applicants respectfully traverse and point out that, on the contrary, Gygi *et al.* never indicate that the correlation between mRNA and protein levels does not exist. Gygi *et al.* only state that the correlation may not be sufficient in **accurately** predicting protein level from the level of the corresponding mRNA transcript (Emphasis added) (see page 1270, Abstract). Contrary to the Examiner's statement, the Gygi data indicate **a general trend** of correlation between protein [expression] and transcript levels (Emphasis added). For example, as shown in Figure 5, the mRNA abundance of **250-300** copies/cell correlates with the protein abundance of **500-1000** x 10<sup>3</sup> copies/cell. The mRNA abundance of **100-200** copies/cell correlates with the protein abundance of **250-500** x 10<sup>3</sup> copies/cell (emphasis added). Therefore, high levels of mRNA **generally** correlate with high levels of proteins. In fact, most data points in Figure 5 did not deviate or scatter away from the general trend of correlation. Thus, the Gygi data, meets the "more likely than not standard" and shows that a positive correlation exists between mRNA and protein. Therefore, Applicants submit that the Examiner's rejection is based on a misrepresentation of the scientific data presented in Gygi *et al.*

In conclusion, the Examiner has not shown that a lack of correlation between gene amplification: polypeptide over-expression, as observed for the *WISP-2* or the *abl* genes, is typical. In fact, contrary to what the Examiner contends, the art indicates that, if a gene is amplified in cancer, it is **more likely than not** that the encoded protein will be expressed at an elevated level. As noted even in Pennica *et al.*, a correlation between DNA amplification: polypeptide over-expression was observed in the case of *WISP-1* and similarly, in Gygi *et al.*, **most genes** showed a correlation between increased mRNA : translated protein. Since the standard is not absolute certainty, a *prima facie* showing of lack of utility has not been made in this instance.



**It is "more likely than not" for amplified genes to have increased mRNA and protein levels**

Applicants submit further exemplary articles to show that, contrary to what the Examiner asserts, the art indicates that, generally, if a gene is amplified in cancer, it is **more likely than not** that the encoded protein will be expressed at an elevated level. For example, Orntoft *et al.* (Mol. and Cell. Proteomics, 2002, Vol.1, pages 37-45) studied transcript levels of 5600 genes in malignant bladder cancers many of which were linked to the gain or loss of chromosomal material using an array-based method. Orntoft *et al.* showed that there was a gene dosage effect and taught that "in general (18 of 23 cases) chromosomal areas with more than 2-fold gain of DNA showed a corresponding increase in mRNA transcripts" (see column 1, abstract). In addition, Hyman *et al.* (Cancer Res., 2002, Vol. 62, pages 6240-45) showed, using CGH analysis and cDNA microarrays which compared DNA copy numbers and mRNA expression of over 12,000 genes in breast cancer tumors and cell lines, that there was "evidence of a prominent global influence of copy number changes on gene expression levels." (see page 6244, column 1, last paragraph). Additional supportive teachings were also provided by Pollack *et al.*, (PNAS, 2002, Vol. 99, pages 12963-12968) who studied a series of primary human breast tumors and showed that "...62% of highly amplified genes show moderately or highly elevated expression, and DNA copy number influences gene expression across a wide range of DNA copy number alterations (deletion, low-, mid- and high-level amplification), and that on average, a 2-fold change in DNA copy number is associated with a corresponding 1.5-fold change in mRNA levels." Thus, these articles collectively teach that in general, gene amplification increases mRNA expression.

In addition, enclosed is a Declaration by Dr. Polakis, principal investigator of the Tumor Antigen Project of Genentech, Inc., the assignee of the present application to show that mRNA expression correlates well with protein levels, in general. As Dr. Polakis explains, the primary focus of the microarray project was to identify tumor cell markers useful as targets for both the diagnosis and treatment of cancer in humans. The scientists working on the project extensively rely on results of microarray experiments in their effort to identify such markers. As Dr. Polakis explains, using microarray analysis, Genentech scientists have identified approximately 200 gene transcripts (mRNAs) that are present in human tumor cells at significantly higher levels than in

corresponding normal human cells. To date, they have generated antibodies that bind to about 30 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human cancer cells and corresponding normal cells. Having compared the levels of mRNA and protein in both the tumor and normal cells analyzed, they found a very good correlation between mRNA and corresponding protein levels. Specifically, in approximately 80% of their observations they have found that increases in the level of a particular mRNA correlates with changes in the level of protein expressed from that mRNA. While the proper legal standard is to show that the existence of correlation between mRNA and polypeptide levels is more likely than not, the showing of approximately 80% correlation for the molecules tested in the Polakis Declaration greatly exceed this legal standard. Based on these experimental data and his vast scientific experience of more than 20 years, Dr. Polakis states that, for human genes, increased mRNA levels typically correlate with an increase in abundance of the encoded protein. He further confirms that "it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein."

Taken together, although there are some examples in the scientific art that do not fit within the central dogma of molecular biology, that there is a correlation between polypeptide and mRNA levels, these instances are exceptions rather than the rule. In the vast majority of amplified genes, the teachings in the art, as exemplified by Orntoft *et al.*, Hyman *et al.*, Pollack *et al.*, and the Polakis declaration, overwhelmingly show that gene amplification influences gene expression at the mRNA and protein levels. Thus, one of skill in the art would reasonably expect in this instance, based on the amplification data for the PRO213-1 gene, that the PRO213-1 protein or a portion thereof, is concomitantly overexpressed. Thus, Applicants submit that the claimed polypeptides including the PRO213-1 proteins have utility in the diagnosis of cancer and based on such a utility, one of skill in the art would know exactly how to use the claimed polypeptides for diagnosis of cancer.

**Even if a *prima facie* case of lack of utility has been established, it should be withdrawn on consideration of the totality of evidence**

Assuming *arguendo* that it is more likely than not that there is no correlation between gene amplification and increased mRNA/protein expression, which Applicants submit is not true, a polypeptide encoded by a gene that is amplified in cancer would **still** have a credible, specific and substantial utility. In support, Applicants submit a Declaration by Avi Ashkenazi, Ph.D., an expert in the field of cancer biology and an inventor of the instant application. Dr. Avi Ashkenazi's Declaration explains that:

even when amplification of a cancer marker gene does not result in significant over-expression of the corresponding gene product, this very absence of gene product over-expression still provides significant information for cancer diagnosis and treatment. Thus, if over-expression of the gene product does not parallel gene amplification in certain tumor types but does so in others, then parallel monitoring of gene amplification and gene product over-expression enables more accurate tumor classification and hence better determination of suitable therapy. In addition, absence of over-expression is crucial information for the practicing clinician. If a gene is amplified but the corresponding gene product is not over-expressed, the clinician accordingly will decide not to treat a patient with agents that target that gene product.

Applicants thus submit that simultaneous testing of gene amplification and gene product over-expression enables more accurate tumor classification, even if the gene-product, the protein, is not over-expressed. This leads to better determination of a suitable therapy. Further, as explained in Dr. Ashkenazi's Declaration, absence of over-expression of the protein itself is crucial information for the practicing clinician. If a gene is amplified in a tumor, but the corresponding gene product is not over-expressed, the clinician will decide not to treat a patient with agents that target that gene product. This not only saves money, but also the patient need not be exposed to the side effects associated with such agents.

This is further supported by the teachings of the attached article by Hanna and Mornin. The article teaches that the HER-2/neu gene has been shown to be amplified and/or over-expressed in 10%-30% of invasive breast cancers and in 40%-60% of intraductal breast carcinoma. Further, the article teaches that diagnosis of breast cancer includes testing both the amplification of the HER-2/neu gene (by FISH) as well as the over-expression of the HER-2/neu

gene product (by IHC). Even when the protein is not over-expressed, the assay relying on both tests leads to a more accurate classification of the cancer and a more effective treatment of it.

Thus, Applicants have demonstrated a credible, specific and substantial asserted utility for the claimed polypeptides, for example, in detecting over-expression or absence of expression of the claimed polypeptides. Further, based on this utility and the disclosure in the specification, one skilled in the art at the time the application was filed would know how to use the claimed polypeptides.

Claims 58-62, and 69-70 stand rejected under 35 U.S.C. §112, first paragraph, because "the specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make or use the invention commensurate in scope with these claims." Specifically, the PTO alleges that "even if there were a patentable use of the protein of SEQ ID 506, variants of 80-99% identity would not be enabled because the specification has not taught one of ordinary skill in the art how to use them or fragments thereof." See page 9 of the instant Office Action.

Without acquiescing to the rejection, Applicants submit that the cancellation of Claims 64-68 renders the rejection of these claims moot. Without acquiescing to the Examiner's position in the current rejections, and without prejudice to further prosecution of the subject-matter in one or more continuation or divisional applications, Claims 58-62 (and, as a consequence, those claims dependent from the same) have been amended to recite "wherein the nucleic acid encoding the polypeptide is amplified in colon or lung tumors" Since the claimed genus is now characterized by a combination of structural and functional features, any person of skill would know how to make and use the invention without undue experimentation based on the general knowledge in the art at the time the invention was made. As the M.P.E.P. states, "The fact that experimentation may be complex does not necessarily make it undue, if the art typically engages in such experimentation" *In re Certain Limited-charge cell Culture Microcarriers*, 221 USPQ 1165, 1174 (Int'l Trade Comm'n 1983), *aff. sub nom.*, *Massachusetts Institute of Technology v A.B. Fortia*, 774 F.2d 1104, 227 USPQ 428 (Fed. Cir. 1985) M.P.E.P. 2164.01.

Accordingly, Applicants submit that pending Claims 58-62 and 69-70 are enabled, as required by 35 U.S.C. §112, first paragraph. The PTO is respectfully requested to reconsider and withdraw the rejection of Claims 58-62, and 69-70 under 35 U.S.C. §112, first paragraph.

In view of the foregoing discussion, Applicants request that the present 35 U.S.C. §101 and §112, first paragraph rejections to the pending claims be withdrawn.

**V. Claim Rejections Under 35 U.S.C. §112, First Paragraph (Enablement) (ATCC Deposit)**

The PTO further alleges that Claims 58-63 and 68-70 stand rejected under 35 U.S.C. §112, first paragraph, because Applicants were not fully compliant with the Budapest Treaty. Specifically, the PTO states that Applicants must state that a viable culture of the deposit would be maintained for 30 years from the date of deposit and for at least five (5) years after the most recent request for the furnishing of a sample of the deposit received by the depository.

Without acquiescing to the rejection, Applicants submit that the cancellation of Claims 64-68 renders the rejection of these claims moot. Further, the sentence beginning on page 378, line 35 has been amended to state, "This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit and for at least five (5) years after the most recent request for the furnishing of a sample of the deposit received by the depository."

Accordingly, Applicants submit that all the requirements of 37 C.F.R. §1.806 are met and that Applicants are fully compliant with the requirements of the Budapest Treaty. Applicants therefore request the PTO to reconsider and withdraw the rejection of the pending claims under 35 U.S.C. §112, first paragraph.

**VI. Claim Rejections – 35 U.S.C. §112, Second paragraph**

Claims 58-70 are rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The Examiner alleges that the claimed polypeptides are not identified as transmembrane proteins, and therefore the term "extracellular" is indefinite.

Without acquiescing to the propriety of this rejection and solely in the interest of expedited prosecution in this case, the term "extracellular domain" is deleted and thus rendering the rejection moot

**VII. Claim Rejections Under 35 U.S.C. §112, First Paragraph (Written Description)**

Claims 58-62, 69, and 70 stand rejected under 35 U.S.C. §112, first paragraph, for allegedly containing subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors had possession of the claimed invention at the time the application was filed. In particular, the PTO notes that "the claims are drawn to polypeptides having at least 80%, 85%, 90%, 95% or 99% sequence identity with a particular disclosed sequence..... The claims do not require that the claimed polypeptides possess any particular biological activity . . . ."

Without acquiescing to the Examiner's position in the current rejections, and without prejudice to further prosecution of the subject-matter in one or more continuation or divisional applications, Claims 58-62 (and, as a consequence, those claims dependent from the same) have been amended to recite "wherein the nucleic acid encoding the polypeptide is amplified in colon or lung tumors."

Thus, this biological activity, coupled with a well defined, and relatively high degree of sequence identity are believed to sufficiently define the claimed genus, such that one skilled in the art would readily recognize that the Applicants were in the possession of the invention claimed at the effective filing date of this application. Hence, the present rejection should be withdrawn.

**VIII. Claim Rejections Under 35 U.S.C. § 102(e)**

Claims 58-69 stand rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Holtzman *et al.* (U.S. Published Patent Application 20020028508) ("Holtzman"), with an effective priority date of April 23, 1998. In particular, the Examiner alleges that Holtzman *et al.* disclose a protein that is 100% identical to the protein of SEQ ID NO: 506.

Claims 58-62, 69 and 70 stand rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Sheppard *et al.* (U.S. Published Patent Application 20030166907) ("Sheppard"), with an effective priority date of June 18, 1997. In particular, the Examiner alleges that Sheppard *et al.* disclose a protein that is 99% identical to the protein of SEQ ID NO: 506.

Without acquiescing to the rejection, Applicants respectfully submit that the cancellation of Claims 64-68 renders the rejection of these claims moot. Claims 58-62 were amended to

recite "the polypeptide of residues 35-273 of SEQ ID NO: 506". In addition, Applicants respectfully submit Declarations under 37 C.F.R. §1.131 by Dr. Goddard, Dr. Godowski, Dr. Gurney, and Dr. Wood, that establish that Applicants had conceived and reduced to practice the invention corresponding to the disclosure of the cited references before June 18 1997, the effective priority date of Holtzman *et al.* and Sheppard *et al.*. The consideration of the Declarations is respectfully requested. Applicants are in process of obtaining signatures of all the inventors of the present application. The Declaration with the signatures of all the inventors will be submitted to the PTO shortly.

**Applicants need to disclose only what is disclosed in the cited reference to support their priority claim**

Applicants respectfully submit that in order to overcome the 35 U.S.C. §102(e) rejection over Holtzman *et al.* and Sheppard *et al.* and support the priority claim, the Declarations by Dr. Goddard, Dr. Godowski, Dr. Gurney, and Dr. Wood ("Declarations") simply need to provide a disclosure commensurate in scope with the disclosure in both Holtzman *et al.* and Sheppard *et al.*

In order to remove a reference as a prior art, "[i]t is sufficient if [the affidavit under Patent Office Rule 131] shows that as much of the claimed invention as is taught in the reference has been reduced to practice by the [patentee] prior to the date of the reference." *In re Stempel*, 241 F.2d 755, 757 (1957). In *In re Stempel*, the patent applicant (Stempel) had claims directed to both (i) a particular genus of chemical compounds (the "generic" claim) and (ii) a single species of chemical compound that was encompassed within that genus (the "species" claim). In support of a rejection under 35 U.S.C. §102, the examiner cited against the application a prior art reference that disclosed the exact chemical compound recited in the "species" claim. In response to the rejection, the patent applicant filed a declaration under 37 C.F.R. §1.131 demonstrating that he had made that specific chemical compound prior to the effective date of the cited prior art reference. The Court found the applicant's 37 C.F.R. § 1.131 declaration effective for swearing behind the cited reference for purposes of both the "species" claim and the "genus" claim. Specifically, the Court stated in support of its decision that "all the applicant can be required to

show is priority with respect to so much of the claimed invention as the reference happens to show. When he has done that he has disposed of the reference.” *Id.* at 759.

Furthermore, the Examiner is respectfully directed to *In re Moore*, 170 USPQ 260 (CCPA 1971), where the holding in *In re Stempel* was affirmed. In *In re Moore*, the patent applicant claimed a particular chemical compound in his patent application and the examiner cited against the applicant a prior art reference under 35 U.S.C. §102 rejection which disclosed the compound but did not disclose any specific utility for the compound. The patent applicant filed a declaration under 37 C.F.R. §1.131 demonstrating that he had made the claimed compound before the effective date of the cited prior art reference, even though he had not yet established a utility for that compound. On appeal, the Court indicated that the 131 declaration filed by the patent applicant was sufficient to remove the cited reference. The Court relied on the established “Stempel Doctrine” to support its decision, stating:

An applicant need **not** be required to show [in a declaration under 37 C.F.R. § 1.131] any more acts with regard to the subject matter claimed that can be carried out by one of ordinary skill in the pertinent art following the description contained in the reference ... the determination of a practical utility when one is not obvious need **not** have been accomplished prior to the date of a reference unless the reference also teaches how to use the compound it describes.

*In re Moore*, 170 USPQ at 267 (emphasis added).

Thus, *In re Moore* confirmed the holding in *In re Stempel* which states that in order to effectively remove a cited reference with a declaration under 37 C.F.R. §1.131, an applicant need only show that portion of his or her claimed invention that appears in the cited reference.

Accordingly, Applicants respectfully submit that the Declarations simply need to show possession of the polypeptide sequence and its encoding polynucleotide sequence and the homology analysis of the polypeptide as disclosed in Holtzman *et al.* and Sheppard *et al.* in order to overcome the 35 U.S.C. §102 rejection over these two references.

As shown in the Declarations, Applicants respectfully submit that Dr. Goddard, Dr. Godowski, Dr. Gurney, and Dr. Wood conceived and reduced to practice the PRO213 polypeptide, which comprises amino acid residues 35-273 of SEQ ID NO:506 as claimed in the present application, and its encoding nucleic acid sequence, in the United States prior to June 18, 1997. The polypeptide encoded by the claimed nucleic acid sequence was also shown



to have homology to human gas 6 protein before the priority date of both prior art patent applications.

As indicated in the Declarations and the brief description of Figure 1 of the present specification, the PRO213 polypeptide is encoded by DNA30943-1163.

Furthermore, as stated in the Declarations, the GSeqEdit database stores cloning, sequencing and functional information for each PRO polypeptide and its encoding nucleic acid sequences according to its DNA number. Copies of the pages from the GSeqEdit database report (with the dates redacted) showing the cloning and sequencing information for the PRO213 polypeptide sequence and its encoding nucleic acid sequence are attached to the Declarations as Exhibit A.

The GSeqEdit report shows the full length nucleic acid sequence for DNA30943-1163 (identified as "DNA30943") and the full length polypeptide sequence encoded by DNA30943-1163. As evidenced from the report and stated in the Declarations, both the nucleic acid and amino acid sequences shown in Exhibit A were obtained prior to June 18, 1997.

In addition, as stated in the Declarations, The DNA sequence of nucleotides 498 to 1216 of the DNA 30943 sequence shown in the GSeqEdit report is identical to that of nucleotides 500-1222 of SEQ ID NO:505 disclosed in the present application. Further, The sequence of amino acid residues 54 to 295 of PRO213 polypeptide shown in the GSeqEdit report is identical to that of amino acids 35-273 of SEQ ID NO: 506 disclosed in the present application. In addition, the report indicates that the polypeptide is homologous to human gas6 protein. Accordingly, the Declarations along with attached Exhibit A clearly show that Applicants were in possession of DNA30943, the polypeptide encoded by DNA30943, and the homology functional information prior to June 18, 1997. Therefore, the Declarations clearly establish that the claimed polypeptides and the nucleic acids encoding thereof, its homology function were conceived and reduced to practice prior to June 18, 1997.

Consequently, based on the holdings of *In re Stempel* and *In re Moore*, Applicants respectfully submit that Holtzman *et al.* and Sheppard *et al.* are not prior art under 102(e) since their priority dates are after the date the instant invention was conceived and reduced to practice in the United States.

Accordingly, the Examiner is respectfully requested to reconsider and withdraw the rejection of Claims 58-69 under 35 U.S.C. §102(e).

**IX. Claim Rejections Under 35 U.S.C. §103(a)**

Claims 70 stands rejected under 35 U.S.C. §103(a) as being unpatentable over Holtzman *et al* in view of Hopp *et al*. For the reasons outlined below, Applicants respectfully disagree with this rejection.

To reject claims in an application under 35 U.S.C. §103, the PTO bears the initial burden of establishing a *prima facie* case of obviousness. *In re Bell*, 26 USPQ2d 1529, 1530 (Fed. Cir. 1993); MPEP § 2142. In order to establish *prima facie* obviousness, three basic criteria must be met.

First, the prior art must provide one of ordinary skill in the art with a suggestion or motivation to modify or combine the teachings of the references relied upon by the PTO to arrive at the claimed invention. Second, the prior art must provide one of ordinary skill in the art with a reasonable expectation of success that the modification or combination suggested by the PTO would succeed. *In re Dow*, 5 USPQ2d 1529, 1531-32 (Fed. Cir. 1988). Third, the prior art, either alone or in combination, must teach or suggest each and every limitation of the rejected claims. *In re Gartside*, 53 USPQ2d 1769 (Fed. Cir. 2000) (Emphasis added). If any one of these criteria are not met, *prima facie* obviousness is not established, and Applicants are not required to show new or unanticipated results. *In re Grabiak*, 226 USPQ 870 (Fed. Cir. 1985).

Applicants submit that the references cited by the PTO are not sufficient to establish a *prima facie* case of obviousness against Claims 70.

As discussed in Section VIII, above, Holtzman *et al* are not prior art, because as evidenced by the Declarations of Dr. Goddard, Dr. Godowski, Dr. Gurney, and Dr. Wood, submitted herewith, the inventors had conceived and reduced the instant invention to practice in the United States before the effective priority date of this reference.

As Holtzman *et al*. are not prior art, the PTO has failed to establish *prima facie* obviousness against Claim 70. Accordingly, Applicants request that the rejection of Claims 70 under 35 U.S.C. § 103(a) as being as being unpatentable over Holtzman *et al*. in view of Hopp *et al*. be withdrawn.

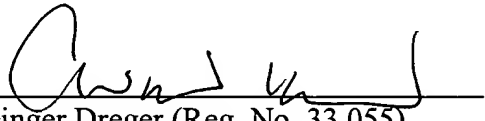
**CONCLUSION**

In conclusion, the present application is believed to be in *prima facie* condition for allowance, and an early action to that effect is respectfully solicited. Should there be any further issues outstanding, the Examiner is invited to contact the undersigned attorney at the telephone number shown below.

Please charge any additional fees, including fees for additional extension of time, or credit overpayment to Deposit Account No. **08-1641** (referencing Attorney's Docket No. **39780-2630 P1C4**).

Respectfully submitted,

Date: October 4, 2004

By:   
Ginger Dreger (Reg. No. 33,055)

**HELLER EHRMAN WHITE & McAULIFFE LLP**

275 Middlefield Road  
Menlo Park, California 94025  
Telephone: (650) 324-7000  
Facsimile: (650) 324-0638

SV 2065656 v1  
10/4/04 11:56 AM (39780.2630)



# TECHNICAL UPDATE

FROM YOUR LABORATORY SERVICES PROVIDER

## HER-2/neu Breast Cancer Predictive Testing

*Julie Sanford Hanna, Ph.D. and Dan Mornin, M.D.*

EACH YEAR, OVER 182,000 WOMEN in the United States are diagnosed with breast cancer, and approximately 45,000 die of the disease.<sup>1</sup> Incidence appears to be increasing in the United States at a rate of roughly 2% per year. The reasons for the increase are unclear, but non-genetic risk factors appear to play a large role.<sup>2</sup>

Five-year survival rates range from approximately 65%-85%, depending on demographic group, with a significant percentage of women experiencing recurrence of their cancer within 10 years of diagnosis. One of the factors most predictive for recurrence once a diagnosis of breast cancer has been made is the number of axillary lymph nodes to which tumor has metastasized. Most node-positive women are given adjuvant therapy, which increases their survival. However, 20%-30% of patients without axillary node involvement also develop recurrent disease, and the difficulty lies in how to identify this high-risk subset of patients. These patients could benefit from increased surveillance, early intervention, and treatment.

Prognostic markers currently used in breast cancer recurrence prediction include tumor size, histological grade, steroid hormone receptor status, DNA ploidy, proliferative index, and cathepsin D status. Expression of growth factor receptors and over-expression of the HER-2/neu oncogene have also been identified as having value regarding treatment regimen and prognosis.

HER-2/neu (also known as c-erbB2) is an oncogene that encodes a transmembrane glycoprotein that is homologous to, but distinct from, the epidermal growth factor receptor. Numerous studies have indicated that high levels of expression of this protein are associated with rapid tumor growth, certain forms of therapy resistance, and shorter disease-free survival. The gene has been shown to be amplified and/or overexpressed in 10%-30% of invasive breast cancers and in 40%-60% of intraductal breast carcinoma.<sup>3</sup>

There are two distinct FDA-approved methods by which HER-2/neu status can be evaluated: immunohistochemistry (IHC, HercepTest™) and FISH (fluorescent in situ hybridization, PathVysion™ Kit). Both methods can be performed on archived and current specimens. The first method allows visual assessment of the amount of HER-2/neu protein present on the cell membrane. The latter method allows direct quantification of the level of gene amplification present in the tumor, enabling differentiation between low- versus high-amplification. At least one study has demonstrated a difference in

recurrence risk in women younger than 40 years of age for low- versus high-amplified tumors (54.5% compared to 85.7%); this is compared to a recurrence rate of 16.7% for patients with no HER-2/neu gene amplification.<sup>4</sup> HER-2/neu status may be particularly important to establish in women with small ( $\leq 1$  cm) tumor size.

The choice of methodology for determination of HER-2/neu status depends in part on the clinical setting. FDA approval for the Vysis FISH test was granted based on clinical trials involving 1549 node-positive patients. Patients received one of three different treatments consisting of different doses of cyclophosphamide, Adriamycin, and 5-fluorouracil (CAF). The study showed that patients with amplified HER-2/neu benefited from treatment with higher doses of adriamycin-based therapy, while those with normal HER-2/neu levels did not. The study therefore identified a sub-set of women, who because they did not benefit from more aggressive treatment, did not need to be exposed to the associated side effects. In addition, other evidence indicates that HER-2/neu amplification in node-negative patients can be used as an independent prognostic indicator for early recurrence, recurrent disease at any time and disease-related death.<sup>5</sup> Demonstration of HER-2/neu gene amplification by FISH has also been shown to be of value in predicting response to chemotherapy in stage-2 breast cancer patients.

Selection of patients for Herceptin® (Trastuzumab) monoclonal antibody therapy, however, is based upon demonstration of HER-2/neu protein overexpression using HercepTest™. Studies using Herceptin® in patients with metastatic breast cancer show an increase in time to disease progression, increased response rate to chemotherapeutic agents and a small increase in overall survival rate. The FISH assays have not yet been approved for this purpose, and studies looking at response to Herceptin® in patients with or without gene amplification status determined by FISH are in progress.

In general, FISH and IHC results correlate well. However, subsets of tumors are found which show discordant results; i.e., protein overexpression without gene amplification or lack of protein overexpression with gene amplification. The clinical significance of such results is unclear. Based on the above considerations, HER-2/neu testing at SHMC/PAML will utilize immunohistochemistry (HercepTest®) as a screen, followed by FISH in IHC-negative cases. Alternatively, either method may be ordered individually depending on the clinical setting or clinician preference.

## CPT code information

## BEST AVAILABLE COPY

### HER-2/neu via IHC

88342 (including interpretive report)

### HER-2/neu via FISH

88271x2 Molecular cytogenetics, DNA probe, each

88274 Molecular cytogenetics, interphase in situ hybridization, analyze 25-99 cells

88291 Cytogenetics and molecular cytogenetics, interpretation and report

## Procedural Information

Immunohistochemistry is performed using the FDA-approved DAKO antibody kit, Herceptest®. The DAKO kit contains reagents required to complete a two-step immunohistochemical staining procedure for routinely processed, paraffin-embedded specimens. Following incubation with the primary rabbit antibody to human HER-2/neu protein, the kit employs a ready-to-use dextran-based visualization reagent. This reagent consists of both secondary goat anti-rabbit antibody molecules with horseradish peroxidase molecules linked to a common dextran polymer backbone, thus eliminating the need for sequential application of link antibody and peroxidase conjugated antibody. Enzymatic conversion of the subsequently added chromogen results in formation of visible reaction product at the antigen site. The specimen is then counterstained; a pathologist using light-microscopy interprets results.

FISH analysis at SHMC/PAML is performed using the FDA-approved PathVysion™ HER-2/neu DNA probe kit, produced by Vysis, Inc. Formalin fixed, paraffin-embedded breast tissue is processed using routine histological methods, and then slides are treated to allow hybridization of DNA probes to the nuclei present in the tissue section. The PathVysion™ kit contains two direct-labeled DNA probes, one specific for the alphoid repetitive DNA (CEP 17, spectrum orange) present at the chromosome 17 centromere and the second for the HER-2/neu oncogene located at 17q11.2-12 (spectrum green). Enumeration of the probes allows a ratio of the number of copies of chromosome 17 to the number of copies of HER-2/neu to be obtained; this enables quantification of low versus high amplification levels, and allows an estimate of the percentage of cells with HER-2/neu gene amplification. The clinically relevant distinction is whether the gene amplification is due to increased gene copy number on the two chromosome 17 homologues normally present or an increase in the number of chromosome 17s in the cells. In the majority of cases, ratio equivalents less than 2.0 are indicative of a normal/negative result, ratios of 2.1 and over indicate that amplification is present and to what degree. Interpretation of this data will be performed and reported from the Vysis-certified Cytogenetics laboratory at SHMC.

1. Wingo, P.A., Tong, T., Bolden, S., "Cancer Statistics", 1995;45:1:8-31.
2. "Cancer Rates and Risks", 4th ed., National Institutes of Health, National Cancer Institute, 1996, p. 120.
3. Slamon, D.J., Clark, G.M., Song, S.G., Levin, W.J., Ulrich, A., McGuire, W.L. "Human breast Cancer: Correlation of relapse and survival with amplification of the her-2/neu oncogene". Science, 235:177-182, 1987.
4. Xing, W.R., Gilchrist, K.W., Harris, C.P., Samson, W., Meisner, L.F. "FISH detection of HER-s/neu oncogene amplification in early onset breast cancer". Breast Cancer Res. And Treatment 39(2):203-212, 1996.
5. Press, M.F., Bernstein, L., Thomas, P.A., Meisner, L.F., Zhou, J.Y., Ma, Y., Hung, G., Robinson, R.A., Harris, C., El-Naggar, A., Slamon, D.J., Phillips, R.N., Ross, J.S., Wolman, S.R., Flom, K.J., "Her-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas", J. Clinical Oncology 15(8):2894-2904, 1997.

*Provided for the clients of*

**PATHOLOGY ASSOCIATES MEDICAL LABORATORIES  
PACLAB NETWORK LABORATORIES  
TRI-CITIES LABORATORY  
TREASURE VALLEY LABORATORY**

*For more information, please contact  
your local representative.*

Impact of DNA Amplification on Gene Expression Patterns in Breast Cancer<sup>1,2</sup>Elizabeth Hyman,<sup>3</sup> Päivikki Kauraniemi,<sup>3</sup> Sampsa Hautaniemi, Maija Wolf, Spyro Mousses, Ester Rozenblum, Markus Ringnér, Guido Sauter, Outi Monni, Abdel Elkahoul, Olli-P. Kallioniemi, and Anne Kallioniemi<sup>4</sup>

Howard Hughes Medical Institute-NIH Research Scholar, Bethesda, Maryland 20892 [E.H.]; Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892 [E.H., P.K., S.H., M.W., S.M., E.R., M.R., A.E., O.K., A.K.]; Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33520 Tampere, Finland [P.K., A.K.]; Signal Processing Laboratory, Tampere University of Technology, FIN-33101 Tampere, Finland [S.H.]; Institute of Pathology, University of Basel, CH-4003 Basel, Switzerland [G.S.]; and Biomedicum Biochip Center, Helsinki University Hospital, Biomedicum Helsinki, FIN-00014 Helsinki, Finland [O.M.]

## ABSTRACT

Genetic changes underlie tumor progression and may lead to cancer-specific expression of critical genes. Over 1100 publications have described the use of comparative genomic hybridization (CGH) to analyze the pattern of copy number alterations in cancer, but very few of the genes affected are known. Here, we performed high-resolution CGH analysis on cDNA microarrays in breast cancer and directly compared copy number and mRNA expression levels of 13,824 genes to quantitate the impact of genomic changes on gene expression. We identified and mapped the boundaries of 24 independent amplicons, ranging in size from 0.2 to 12 Mb. Throughout the genome, both high- and low-level copy number changes had a substantial impact on gene expression, with 44% of the highly amplified genes showing overexpression and 10.5% of the highly overexpressed genes being amplified. Statistical analysis with random permutation tests identified 270 genes whose expression levels across 14 samples were systematically attributable to gene amplification. These included most previously described amplified genes in breast cancer and many novel targets for genomic alterations, including the *HOXB7* gene, the presence of which in a novel amplicon at 17q21.3 was validated in 10.2% of primary breast cancers and associated with poor patient prognosis. In conclusion, CGH on cDNA microarrays revealed hundreds of novel genes whose overexpression is attributable to gene amplification. These genes may provide insights to the clonal evolution and progression of breast cancer and highlight promising therapeutic targets.

## INTRODUCTION

Gene expression patterns revealed by cDNA microarrays have facilitated classification of cancers into biologically distinct categories, some of which may explain the clinical behavior of the tumors (1-6). Despite this progress in diagnostic classification, the molecular mechanisms underlying gene expression patterns in cancer have remained elusive, and the utility of gene expression profiling in the identification of specific therapeutic targets remains limited.

Accumulation of genetic defects is thought to underlie the clonal evolution of cancer. Identification of the genes that mediate the effects of genetic changes may be important by highlighting transcripts that are actively involved in tumor progression. Such transcripts and their encoded proteins would be ideal targets for anticancer therapies, as demonstrated by the clinical success of new therapies against amplified oncogenes, such as *ERBB2* and *EGFR* (7, 8), in breast cancer and other solid tumors. Besides amplifications of known oncogenes, over

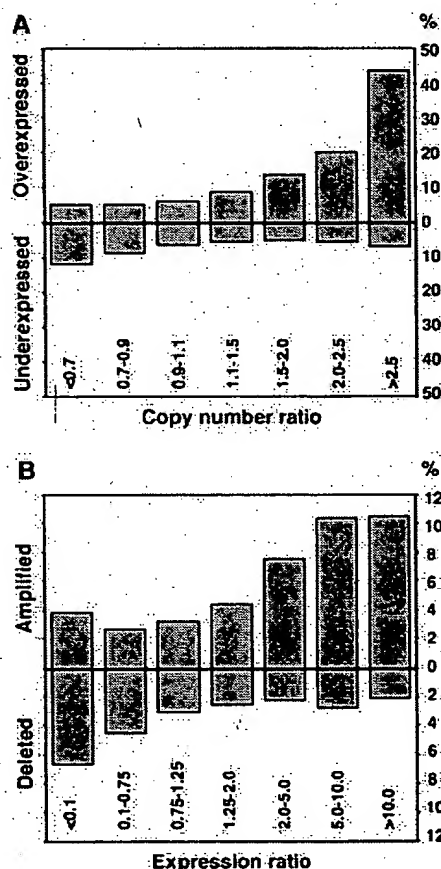


Fig. 1. Impact of gene copy number on global gene expression levels. A, percentage of over- and underexpressed genes (Y axis) according to copy number ratios (X axis). Threshold values used for over- and underexpression were  $>2.184$  (global upper 7% of the cDNA ratios) and  $<0.4826$  (global lower 7% of the expression ratios). B, percentage of amplified and deleted genes according to expression ratios. Threshold values for amplification and deletion were  $>1.5$  and  $<0.7$ .

20 recurrent regions of DNA amplification have been mapped in breast cancer by CGH<sup>5</sup> (9, 10). However, these amplicons are often large and poorly defined, and their impact on gene expression remains unknown.

We hypothesized that genome-wide identification of those gene expression changes that are attributable to underlying gene copy number alterations would highlight transcripts that are actively involved in the causation or maintenance of the malignant phenotype. To identify such transcripts, we applied a combination of cDNA and CGH microarrays to: (a) determine the global impact that gene copy number variation plays in breast cancer development and progression; and (b) identify and characterize those genes whose mRNA expres-

Received 5/29/02; accepted 8/28/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> Supported in part by the Academy of Finland, Emil Aaltonen Foundation, the Finnish Cancer Society, the Pirkanmaa Cancer Society, the Pirkanmaa Cultural Foundation, the Finnish Breast Cancer Group, the Foundation for the Development of Laboratory Medicine, the Medical Research Fund of the Tampere University Hospital, the Foundation for Commercial and Technical Sciences, and the Swedish Research Council.

<sup>2</sup> Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org>).

<sup>3</sup> Contributed equally to this work.

<sup>4</sup> To whom requests for reprints should be addressed, at Laboratory of Cancer Genetics, Institute of Medical Technology, Lenkkilijankatu 6, FIN-33520 Tampere, Finland. Phone: 358-3247-4125; Fax: 358-3247-4168; E-mail: anne.kallioniemi@uta.fi.

<sup>5</sup> The abbreviations used are: CGH, comparative genomic hybridization; FISH, fluorescence *in situ* hybridization; RT-PCR, reverse transcription-PCR.

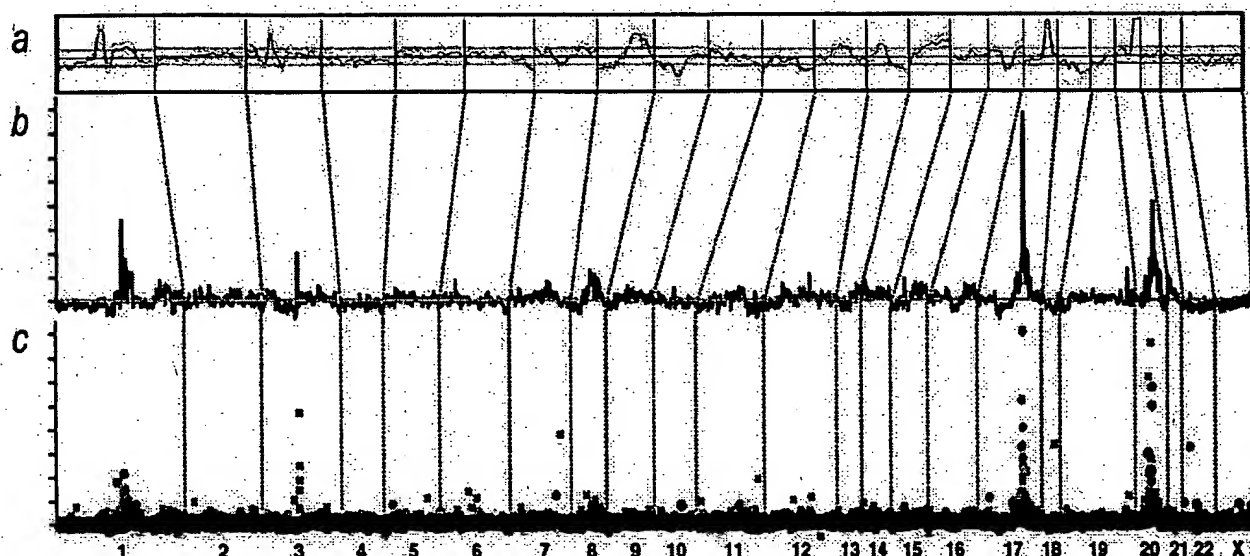


Fig. 2. Genome-wide copy number and expression analysis in the MCF-7 breast cancer cell line. *A*, chromosomal CGH analysis of MCF-7. The copy number ratio profile (blue line) across the entire genome from 1p telomere to Xq telomere is shown along with  $\pm 1$  SD (orange lines). The black horizontal line indicates a ratio of 1.0; red line, a ratio of 0.8; and green line, a ratio of 1.2. *B–C*, genome-wide copy number analysis in MCF-7 by CGH on cDNA microarray. The copy number ratios were plotted as a function of the position of the cDNA clones along the human genome. In *B*, individual data points are connected with a line, and a moving median of 10 adjacent clones is shown. Red horizontal line, the copy number ratio of 1.0. In *C*, individual data points are labeled by color coding according to cDNA expression ratios. The bright red dots indicate the upper 2%, and dark red dots, the next 5% of the expression ratios in MCF-7 cells (overexpressed genes); bright green dots indicate the lowest 2%, and dark green dots, the next 5% of the expression ratios (underexpressed genes); the rest of the observations are shown with black crosses. The chromosome numbers are shown at the bottom of the figure, and chromosome boundaries are indicated with a dashed line.

sion is most significantly associated with amplification of the corresponding genomic template.

## MATERIALS AND METHODS

**Breast Cancer Cell Lines.** Fourteen breast cancer cell lines (BT-20, BT-474, HCC1428, Hs578t, MCF7, MDA-361, MDA-436, MDA-453, MDA-468, SKBR-3, T-47D, UACC812, ZR-75-1, and ZR-75-30) were obtained from the American Type Culture Collection (Manassas, VA). Cells were grown under recommended culture conditions. Genomic DNA and mRNA were isolated using standard protocols.

**Copy Number and Expression Analyses by cDNA Microarrays.** The preparation and printing of the 13,824 cDNA clones on glass slides were performed as described (11–13). Of these clones, 244 represented uncharacterized expressed sequence tags, and the remainder corresponded to known genes. CGH experiments on cDNA microarrays were done as described (14, 15). Briefly, 20  $\mu$ g of genomic DNA from breast cancer cell lines and normal human WBCs were digested for 14–18 h with *AhaI* and *RsaI* (Life Technologies, Inc., Rockville, MD) and purified by phenol/chloroform extraction. Six  $\mu$ g of digested cell line DNAs were labeled with Cy3-dUTP (Amersham Pharmacia) and normal DNA with Cy5-dUTP (Amersham Pharmacia) using the Bioprime Labeling kit (Life Technologies, Inc.). Hybridization (14, 15) and posthybridization washes (13) were done as described. For the expression analyses, a standard reference (Universal Human Reference RNA; Stratagene, La Jolla, CA) was used in all experiments. Forty  $\mu$ g of reference RNA were labeled with Cy3-dUTP and 3.5  $\mu$ g of test mRNA with Cy5-dUTP, and the labeled cDNAs were hybridized on microarrays as described (13, 15). For both microarray analyses, a laser confocal scanner (Agilent Technologies, Palo Alto, CA) was used to measure the fluorescence intensities at the target locations using the DEARRAY software (16). After background subtraction, average intensities at each clone in the test hybridization were divided by the average intensity of the corresponding clone in the control hybridization. For the copy number analysis, the ratios were normalized on the basis of the distribution of ratios of all targets on the array and for the expression analysis on the basis of 88 housekeeping genes, which were spotted four times onto the array. Low quality measurements (*i.e.*, copy number data with mean reference intensity <100 fluorescent units, and expression data with both test and reference intensity <100 fluorescent units and/or with spot size <50 units)

were excluded from the analysis and were treated as missing values. The distributions of fluorescence ratios were used to define cutpoints for increased/decreased copy number. Genes with CGH ratio >1.43 (representing the upper 5% of the CGH ratios across all experiments) were considered to be amplified, and genes with ratio <0.73 (representing the lower 5%) were considered to be deleted.

**Statistical Analysis of CGH and cDNA Microarray Data.** To evaluate the influence of copy number alterations on gene expression, we applied the following statistical approach. CGH and cDNA calibrated intensity ratios were log-transformed and normalized using median centering of the values in each cell line. Furthermore, cDNA ratios for each gene across all 14 cell lines were median centered. For each gene, the CGH data were represented by a vector that was labeled 1 for amplification (ratio, >1.43) and 0 for no amplification. Amplification was correlated with gene expression using the signal-to-noise statistics (1). We calculated a weight,  $w_g$ , for each gene as follows:

$$w_g = \frac{m_{g1} - m_{g0}}{\sigma_{g1} + \sigma_{g0}}$$

where  $m_{g1}$ ,  $\sigma_{g1}$ , and  $m_{g0}$ ,  $\sigma_{g0}$  denote the means and SDs for the expression levels for amplified and nonamplified cell lines, respectively. To assess the statistical significance of each weight, we performed 10,000 random permutations of the label vector. The probability that a gene had a larger or equal weight by random permutation than the original weight was denoted by  $\alpha$ . A low  $\alpha$  (<0.05) indicates a strong association between gene expression and amplification.

**Genomic Localization of cDNA Clones and Amplicon Mapping.** Each cDNA clone on the microarray was assigned to a Unigene cluster using the Unigene Build 141.<sup>6</sup> A database of genomic sequence alignment information for mRNA sequences was created from the August 2001 freeze of the University of California Santa Cruz's GoldenPath database.<sup>7</sup> The chromosome and bp positions for each cDNA clone were then retrieved by relating these data sets. Amplicons were defined as a CGH copy number ratio >2.0 in at least two adjacent clones in two or more cell lines or a CGH ratio >2.0 in at least three adjacent clones in a single cell line. The amplicon start and end positions were

<sup>6</sup> Internet address: [http://research.nhgri.nih.gov/microarray/downloadable\\_cdna.html](http://research.nhgri.nih.gov/microarray/downloadable_cdna.html).

<sup>7</sup> Internet address: [www.genome.ucsc.edu](http://www.genome.ucsc.edu).

Table 1 Summary of independent amplicons in 14 breast cancer cell lines by CGH microarray

Location	Start (Mb)	End (Mb)	Size (Mb)
1p13	132.79	132.94	0.2
1q21	173.92	177.25	3.3
1q22	179.28	179.57	0.3
3p14	71.94	74.66	2.7
7p12.1-7p11.2	55.62	60.95	5.3
7q31	125.73	130.96	5.2
7q32	140.01	140.68	0.7
8q21.11-8q21.13	86.45	92.46	6.0
8q21.3	98.45	103.05	4.6
8q23.3-8q24.14	129.88	142.15	12.3
8q24.22	151.21	152.16	1.0
9p13	38.65	39.25	0.6
13q22-q31	77.15	81.38	4.2
16q22	86.70	87.62	0.9
17q11	29.30	30.85	1.6
17q12-q21.2	39.79	42.80	3.0
17q21.32-q21.33	52.47	55.80	3.3
17q22-q23.3	63.81	69.70	5.9
17q23.3-q24.3	69.93	74.99	5.1
19q13	40.63	41.40	0.8
20q11.22	34.59	35.85	1.3
20q13.12	44.00	45.62	1.6
20q13.12-q13.13	46.45	49.43	3.0
20q13.2-q13.32	51.32	59.12	7.8

extended to include neighboring nonamplified clones (ratio, <1.5). The amplicon size determination was partially dependent on local clone density.

**FISH.** Dual-color interphase FISH to breast cancer cell lines was done as described (17). Bacterial artificial chromosome clone RP11-361K8 was labeled with SpectrumOrange (Vysis, Downers Grove, IL), and SpectrumOrange-labeled probe for *EGFR* was obtained from Vysis. SpectrumGreen-labeled chromosome 7 and 17 centromere probes (Vysis) were used as a reference. A tissue microarray containing 612 formalin-fixed, paraffin-embedded primary breast cancers (17) was applied in FISH analyses as described (18). The use of these specimens was approved by the Ethics Committee of the University of Basel and by the NIH. Specimens containing a 2-fold or higher increase in the number of test probe signals, as compared with corresponding centromere signals, in at least 10% of the tumor cells were considered to be amplified. Survival analysis was performed using the Kaplan-Meier method and the log-rank test.

**RT-PCR.** The *HOXB7* expression level was determined relative to *GAPDH*. Reverse transcription and PCR amplification were performed using Access RT-PCR System (Promega Corp., Madison, WI) with 10 ng of mRNA as a template. *HOXB7* primers were 5'-GAGCAGAGGACTCGGACTT-3' and 5'-GCGTCAGGTAGCGATTGTAG-3'.

## RESULTS

**Global Effect of Copy Number on Gene Expression.** 13,824 arrayed cDNA clones were applied for analysis of gene expression and gene copy number (CGH microarrays) in 14 breast cancer cell lines. The results illustrate a considerable influence of copy number on gene expression patterns. Up to 44% of the highly amplified transcripts (CGH ratio, >2.5) were overexpressed (i.e., belonged to the global upper 7% of expression ratios), compared with only 6% for genes with normal copy number levels (Fig. 1A). Conversely, 10.5% of the transcripts with high-level expression (cDNA ratio, >10) showed increased copy number (Fig. 1B). Low-level copy number increases and decreases were also associated with similar, although less dramatic, outcomes on gene expression (Fig. 1).

**Identification of Distinct Breast Cancer Amplicons.** Base-pair locations obtained for 11,994 cDNAs (86.8%) were used to plot copy number changes as a function of genomic position (Fig. 2, Supplement Fig. A). The average spacing of clones throughout the genome was 267 kb. This high-resolution mapping identified 24 independent breast cancer amplicons, spanning from 0.2 to 12 Mb of DNA (Table 1). Several amplification sites detected previously by chromosomal

CGH were validated, with 1q21, 17q12-q21.2, 17q22-q23, 20q13.1, and 20q13.2 regions being most commonly amplified. Furthermore, the boundaries of these amplicons were precisely delineated. In addition, novel amplicons were identified at 9p13 (38.65-39.25 Mb), and 17q21.3 (52.47-55.80 Mb).

**Direct Identification of Putative Amplification Target Genes.** The cDNA/CGH microarray technique enables the direct correlation of copy number and expression data on a gene-by-gene basis throughout the genome. We directly annotated high-resolution CGH plots with gene expression data using color coding. Fig. 2C shows that most of the amplified genes in the MCF-7 breast cancer cell line at 1p13, 17q22-q23, and 20q13 were highly overexpressed. A view of chromosome 7 in the MDA-468 cell line implicates *EGFR* as the most highly overexpressed and amplified gene at 7p11-p12 (Fig. 3A). In BT-474, the two known amplicons at 17q12 and 17q22-q23 contained numerous highly overexpressed genes (Fig. 3B). In addition, several genes, including the homeobox genes *HOXB2* and *HOXB7*, were highly amplified in a previously undescribed independent amplicon at 17q21.3. *HOXB7* was systematically amplified (as validated by FISH, Fig. 3B, inset) as well as overexpressed (as verified by RT-PCR, data not shown) in BT-474, UACC812, and ZR-75-30 cells. Furthermore, this novel

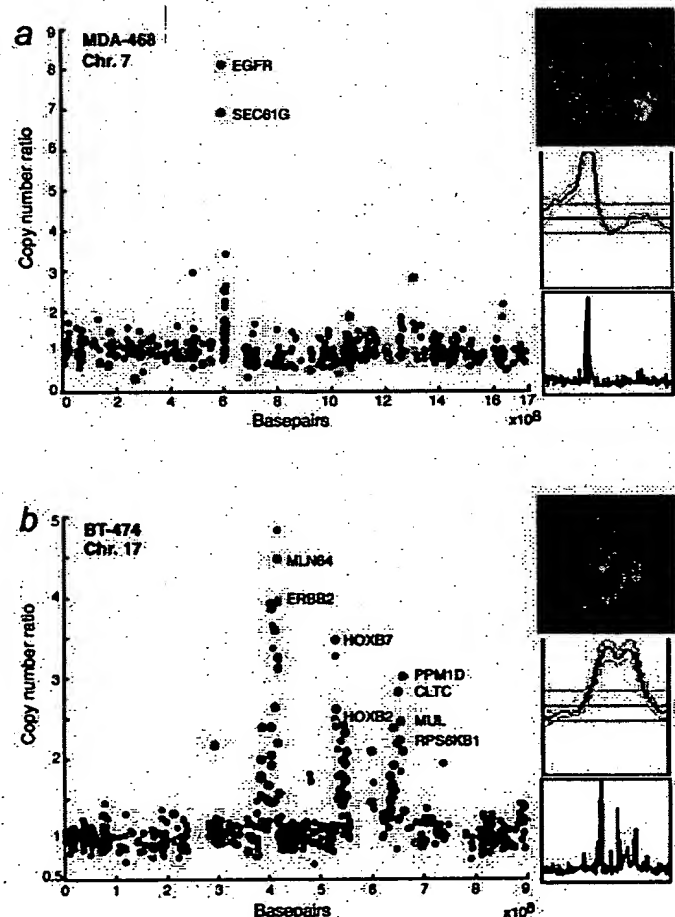
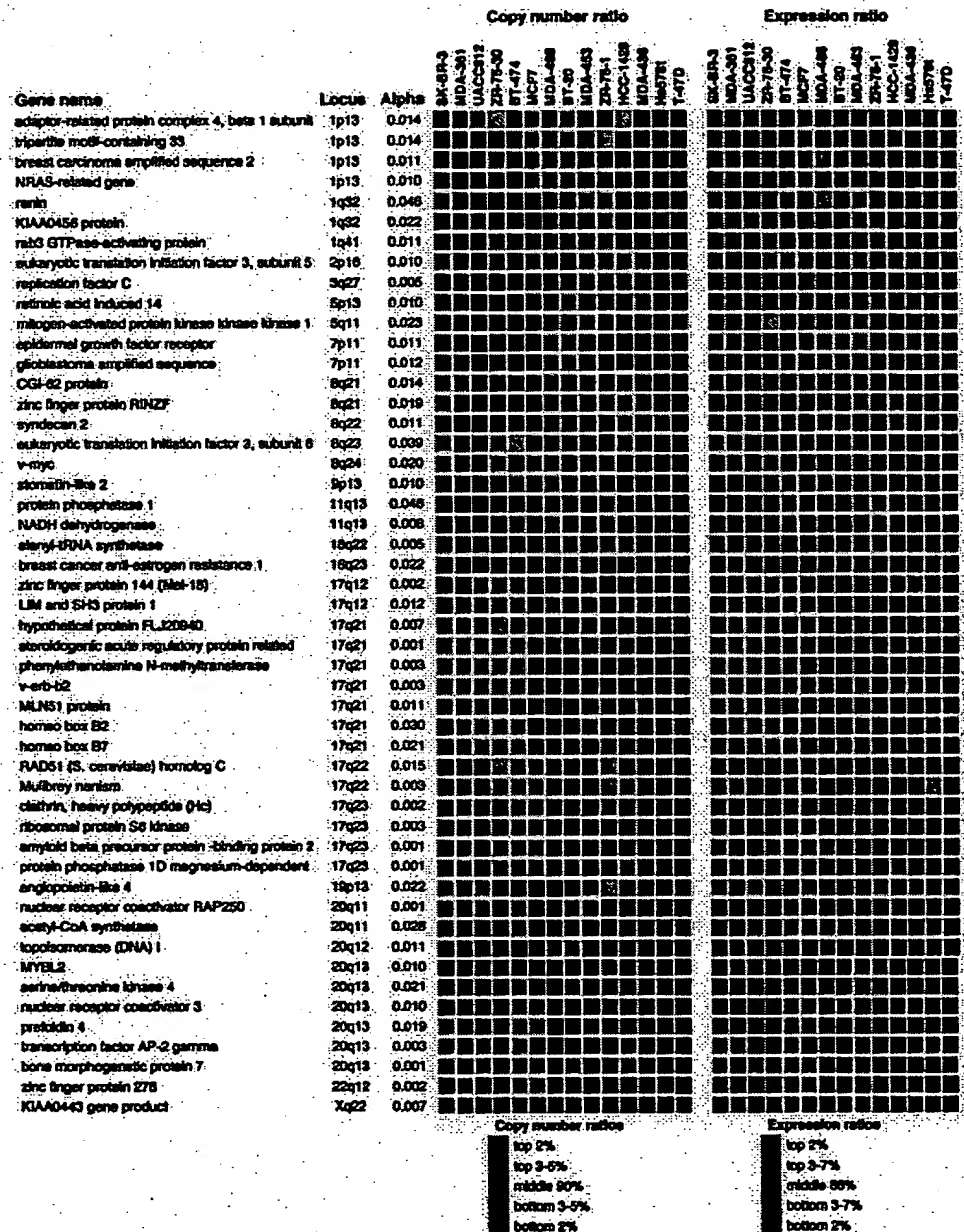


Fig. 3. Annotation of gene expression data on CGH microarray profiles. A, genes in the 7p11-p12 amplicon in the MDA-468 cell line are highly expressed (red dots) and include the *EGFR* oncogene. B, several genes in the 17q12, 17q21.3, and 17q23 amplicons in the BT-474 breast cancer cell line are highly overexpressed (red) and include the *HOXB7* gene. The data labels and color coding are as indicated for Fig. 2C. Insets show chromosomal CGH profiles for the corresponding chromosomes and validation of the increased copy number by interphase FISH using *EGFR* (red) and chromosome 7 centromere probe (green) to MDA-468 (A) and *HOXB7*-specific probe (red) and chromosome 17 centromere (green) to BT-474 cells (B).



Fig. 4. List of 50 genes with a statistically significant correlation ( $\alpha$  value  $<0.05$ ) between gene copy number and gene expression. Name, chromosomal location, and the  $\alpha$  value for each gene are indicated. The genes have been ordered according to their position in the genome. The color maps on the right illustrate the copy number and expression ratio patterns in the 14 cell lines. The key to the color code is shown at the bottom of the graph. Gray squares, missing values. The complete list of 270 genes is shown in supplemental Fig. B.



amplification was validated to be present in 10.2% of 363 primary breast cancers by FISH to a tissue microarray and was associated with poor prognosis of the patients ( $P = 0.001$ ).

**Statistical Identification and Characterization of 270 Highly Expressed Genes in Amplicons.** Statistical comparison of expression levels of all genes as a function of gene amplification identified 270 genes whose expression was significantly influenced by copy number across all 14 cell lines (Fig. 4, Supplemental Fig. B). According to the gene ontology data,<sup>8</sup> 91 of the 270 genes represented hypothetical proteins or genes with no functional annotation, whereas 179 had associated functional information available. Of these, 151 (84%) are implicated in apoptosis, cell proliferation, signal transduction, and transcription, whereas 28 (16%) had functional annotations that could not be directly linked with cancer.

## DISCUSSION

The importance of recurrent gene and chromosome copy number changes in the development and progression of solid tumors has been characterized in >1000 publications applying CGH<sup>9</sup> (9, 10), as well as in a large number of other molecular cytogenetic, cytogenetic, and molecular genetic studies. The effects of these somatic genetic changes on gene expression levels have remained largely unknown, although a few studies have explored gene expression changes occurring in specific amplicons (15, 19–21). Here, we applied genome-wide cDNA microarrays to identify transcripts whose expression changes were attributable to underlying gene copy number alterations in breast cancer.

The overall impact of copy number on gene expression patterns was substantial with the most dramatic effects seen in the case of high-

<sup>8</sup> Internet address: <http://www.geneontology.org/>.

<sup>9</sup> Internet address: <http://www.ncbi.nlm.nih.gov/entrez>.

level copy number increase. Low-level copy number gains and losses also had a significant influence on expression levels of genes in the regions affected, but these effects were more subtle on a gene-by-gene basis than those of high-level amplifications. However, the impact of low-level gains on the dysregulation of gene expression patterns in cancer may be equally important if not more important than that of high-level amplifications. Aneuploidy and low-level gains and losses of chromosomal arms represent the most common types of genetic alterations in breast and other cancers and, therefore, have an influence on many genes. Our results in breast cancer extend the recent studies on the impact of aneuploidy on global gene expression patterns in yeast cells, acute myeloid leukemia, and a prostate cancer model system (22–24).

The CGH microarray analysis identified 24 independent breast cancer amplicons. We defined the precise boundaries for many amplicons detected previously by chromosomal CGH (9, 10, 25, 26) and also discovered novel amplicons that had not been detected previously, presumably because of their small size (only 1–2 Mb) or close proximity to other larger amplicons. One of these novel amplicons involved the homeobox gene region at 17q21.3 and led to the overexpression of the *HOXB7* and *HOXB2* genes. The homeodomain transcription factors are known to be key regulators of embryonic development and have been occasionally reported to undergo aberrant expression in cancer (27, 28). *HOXB7* transfection induced cell proliferation in melanoma, breast, and ovarian cancer cells and increased tumorigenicity and angiogenesis in breast cancer (29–32). The present results imply that gene amplification may be a prominent mechanism for overexpressing *HOXB7* in breast cancer and suggest that *HOXB7* contributes to tumor progression and confers an aggressive disease phenotype in breast cancer. This view is supported by our finding of amplification of *HOXB7* in 10% of 363 primary breast cancers, as well as an association of amplification with poor prognosis of the patients.

We carried out a systematic search to identify genes whose expression levels across all 14 cell lines were attributable to amplification status. Statistical analysis revealed 270 such genes (representing ~2% of all genes on the array), including not only previously described amplified genes, such as *HER-2*, *MYC*, *EGFR*, ribosomal protein s6 kinase, and *AIB3*, but also numerous novel genes such as *NRAS-related gene* (1p13), *syndecan-2* (8q22), and *bone morphogenic protein* (20q13.1), whose activation by amplification may similarly promote breast cancer progression. Most of the 270 genes have not been implicated previously in breast cancer development and suggest novel pathogenetic mechanisms. Although we would not expect all of them to be causally involved, it is intriguing that 84% of the genes with associated functional information were implicated in apoptosis, cell proliferation, signal transduction, transcription, or other cellular processes that could directly imply a possible role in cancer progression. Therefore, a detailed characterization of these genes may provide biological insights to breast cancer progression and might lead to the development of novel therapeutic strategies.

In summary, we demonstrate application of cDNA microarrays to the analysis of both copy number and expression levels of over 12,000 transcripts throughout the breast cancer genome, roughly once every 267 kb. This analysis provided: (a) evidence of a prominent global influence of copy number changes on gene expression levels; (b) a high-resolution map of 24 independent amplicons in breast cancer; and (c) identification of a set of 270 genes, the overexpression of which was statistically attributable to gene amplification. Characterization of a novel amplicon at 17q21.3 implicated amplification and overexpression of the *HOXB7* gene in breast cancer, including a clinical association

between *HOXB7* amplification and poor patient prognosis. Overall, our results illustrate how the identification of genes activated by gene amplification provides a powerful approach to highlight genes with an important role in cancer as well as to prioritize and validate putative targets for therapy development.

## REFERENCES

- Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* (Wash. DC), 286: 531–537, 1999.
- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* (Lond.), 403: 503–511, 2000.
- Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* (Lond.), 406: 536–540, 2000.
- Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., et al. Molecular portraits of human breast tumours. *Nature* (Lond.), 406: 747–752, 2000.
- Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. Delineation of prognostic biomarkers in prostate cancer. *Nature* (Lond.), 412: 822–826, 2001.
- Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA*, 98: 10869–10874, 2001.
- Ross, J. S., and Fletcher, J. A. The *HER-2/neu* oncogene: prognostic factor, predictive factor and target for therapy. *Semin. Cancer Biol.*, 9: 125–138, 1999.
- Arteaga, C. L. The epidermal growth factor receptor: from mutant oncogene in nonhuman cancers to therapeutic target in human neoplasia. *J. Clin. Oncol.*, 19: 32–40, 2001.
- Knuutila, S., Björkqvist, A. M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M. L., Tapper, J., Pere, H., El-Rifai, W., et al. DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. *Am. J. Pathol.*, 152: 1107–1123, 1998.
- Knuutila, S., Autio, K., and Aalto, Y. Online access to CGH data of DNA sequence copy number changes. *Am. J. Pathol.*, 157: 689, 2000.
- DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.*, 14: 457–460, 1996.
- Shalon, D., Smith, S. J., and Brown, P. O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.*, 6: 639–645, 1996.
- Mousses, S., Bittner, M. L., Chen, Y., Dougherty, E. R., Baxevanis, A., Meltzer, P. S., and Trent, J. M. Gene expression analysis by cDNA microarrays. In: F. J. Livesey and S. P. Hunt (eds.), *Functional Genomics*, pp. 113–137. Oxford: Oxford University Press, 2000.
- Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.*, 23: 41–46, 1999.
- Monni, O., Bärilund, M., Mousses, S., Kononen, J., Sauter, G., Heiskanen, M., Paavola, P., Avela, K., Chen, Y., Bittner, M. L., and Kallioniemi, A. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. *Proc. Natl. Acad. Sci. USA*, 98: 5711–5716, 2001.
- Chen, Y., Dougherty, E. R., and Bittner, M. L. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics*, 2: 364–374, 1997.
- Bärilund, M., Forozan, F., Kononen, J., Bubendorf, L., Chen, Y., Bittner, M. L., Thorhorst, J., Haas, P., Bucher, C., Sauter, G., et al. Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. *J. Natl. Cancer Inst.*, 92: 1252–1259, 2000.
- Andersen, C. L., Hostetter, G., Grigoryan, A., Sauter, G., and Kallioniemi, A. Improved procedure for fluorescence *in situ* hybridization on tissue microarrays. *Cytometry*, 45: 83–86, 2001.
- Kauraniemi, P., Bärilund, M., Monni, O., and Kallioniemi, A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. *Cancer Res.*, 61: 8235–8240, 2001.
- Clark, J., Edwards, S., John, M., Flohr, P., Gordon, T., Maillard, K., Giddings, I., Brown, C., Bagherzadeh, A., Campbell, C., Shipley, J., Wooster, R., and Cooper, C. S. Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. *Genes Chromosomes Cancer*, 34: 104–114, 2002.
- Varis, A., Wolf, M., Monni, O., Vakkari, M. L., Kokkola, A., Moskaluk, C., Frierson, H., Powell, S. M., Knuutila, S., Kallioniemi, A., and El-Rifai, W. Targets of gene amplification and overexpression at 17q in gastric cancer. *Cancer Res.*, 62: 2625–2629, 2002.
- Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., Friend, S. H., and Marton, M. J.

- Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat. Genet.*, 25: 333-337, 2000.
23. Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de La Chapelle, A., and Krahe, R. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc. Natl. Acad. Sci. USA*, 98: 1124-1129, 2001.
  24. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., Linchan, W. M., Cunha, G. R., and Ried, T. The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. *Cancer Res.*, 61: 8143-8149, 2001.
  25. Bärklund, M., Tirkkonen, M., Forozan, F., Tanner, M. M., Kallioniemi, O. P., and Kallioniemi, A. Increased copy number at 17q22-q24 by CGH in breast cancer is due to high-level amplification of two separate regions. *Genes Chromosomes Cancer*, 20: 372-376, 1997.
  26. Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Isola, J., Kuukasjärvi, T., Collins, C., Kowbel, D., Guan, X. Y., Trent, J., Gray, J. W., Meltzer, P., and Kallioniemi O. P. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. *Cancer Res.*, 56: 3441-3445, 1996.
  27. Cillo, C., Faiella, A., Cantile, M., and Boncinelli, E. Homeobox genes and cancer. *Exp. Cell Res.*, 248: 1-9, 1999.
  28. Cillo, C., Cantile, M., Faiella, A., and Boncinelli, E. Homeobox genes in normal and malignant cells. *J. Cell. Physiol.*, 188: 161-169, 2001.
  29. Care, A., Silvani, A., Meccia, E., Mattia, G., Stoppacciaro, A., Parmiani, G., Peschle, C., and Colombo, M. P. HOXB7 constitutively activates basic fibroblast growth factor in melanomas. *Mol. Cell. Biol.*, 16: 4842-4851, 1996.
  30. Care, A., Silvani, A., Meccia, E., Mattia, G., Peschle, C., and Colombo, M. P. Transduction of the SkBr3 breast carcinoma cell line with the HOXB7 gene induces bFGF expression, increases cell proliferation and reduces growth factor dependence. *Oncogene*, 16: 3285-3289, 1998.
  31. Care, A., Felicetti, F., Meccia, E., Bottero, L., Parenza, M., Stoppacciaro, A., Peschle, C., and Colombo, M. P. HOXB7: a key factor for tumor-associated angiogenic switch. *Cancer Res.*, 61: 6532-6539, 2001.
  32. Naora, H., Yang, Y. Q., Montz, F. J., Seidman, J. D., Kurman, R. J., and Roden, R. B. A serologically identified tumor antigen encoded by a homeobox gene promotes growth of ovarian epithelial cells. *Proc. Natl. Acad. Sci. USA*, 98: 4060-4065, 2001.

# Genome-wide Study of Gene Copy Numbers, Transcripts, and Protein Levels in Pairs of Non-invasive and Invasive Human Transitional Cell Carcinomas\*

Torben F. Ørntoft‡§, Thomas Thykjaer||, Frederic M. Waldman||, Hans Wolf\*\*, and Julio E. Celis‡‡

Gain and loss of chromosomal material is characteristic of bladder cancer, as well as malignant transformation in general. The consequences of these changes at both the transcription and translation levels is at present unknown partly because of technical limitations. Here we have attempted to address this question in pairs of non-invasive and invasive human bladder tumors using a combination of technology that included comparative genomic hybridization, high density oligonucleotide array-based monitoring of transcript levels (5600 genes), and high resolution two-dimensional gel electrophoresis. The results showed that there is a gene dosage effect that in some cases superimposes on other regulatory mechanisms. This effect depended ( $p < 0.015$ ) on the magnitude of the comparative genomic hybridization change. In general (18 of 23 cases), chromosomal areas with more than 2-fold gain of DNA showed a corresponding increase in mRNA transcripts. Areas with loss of DNA, on the other hand, showed either reduced or unaltered transcript levels. Because most proteins resolved by two-dimensional gels are unknown it was only possible to compare mRNA and protein alterations in relatively few cases of well focused abundant proteins. With few exceptions we found a good correlation ( $p < 0.005$ ) between transcript alterations and protein levels. The implications, as well as limitations, of the approach are discussed. *Molecular & Cellular Proteomics* 1:37–45, 2002.

Aneuploidy is a common feature of most human cancers (1), but little is known about the genome-wide effect of this

phenomenon at both the transcription and translation levels. High throughput array studies of the breast cancer cell line BT474 has suggested that there is a correlation between DNA copy numbers and gene expression in highly amplified areas (2), and studies of individual genes in solid tumors have revealed a good correlation between gene dose and mRNA or protein levels in the case of c-erb-B2, cyclin d1, *ems1*, and N-myc (3–5). However, a high cyclin D1 protein expression has been observed without simultaneous amplification (4), and a low level of c-myc copy number increase was observed without concomitant c-myc protein overexpression (6).

In human bladder tumors, karyotyping, fluorescent *in situ* hybridization, and comparative genomic hybridization (CGH)<sup>1</sup> have revealed chromosomal aberrations that seem to be characteristic of certain stages of disease progression. In the case of non-invasive pTa transitional cell carcinomas (TCCs), this includes loss of chromosome 9 or parts of it, as well as loss of Y in males. In minimally invasive pT1 TCCs, the following alterations have been reported: 2q–, 11p–, 1q+, 11q13+, 17q+, and 20q+ (7–12). It has been suggested that these regions harbor tumor suppressor genes and oncogenes; however, the large chromosomal areas involved often contain many genes, making meaningful predictions of the functional consequences of losses and gains very difficult.

In this investigation we have combined genome-wide technology for detecting genomic gains and losses (CGH) with gene expression profiling techniques (microarrays and proteomics) to determine the effect of gene copy number on transcript and protein levels in pairs of non-invasive and invasive human bladder TCCs.

## EXPERIMENTAL PROCEDURES

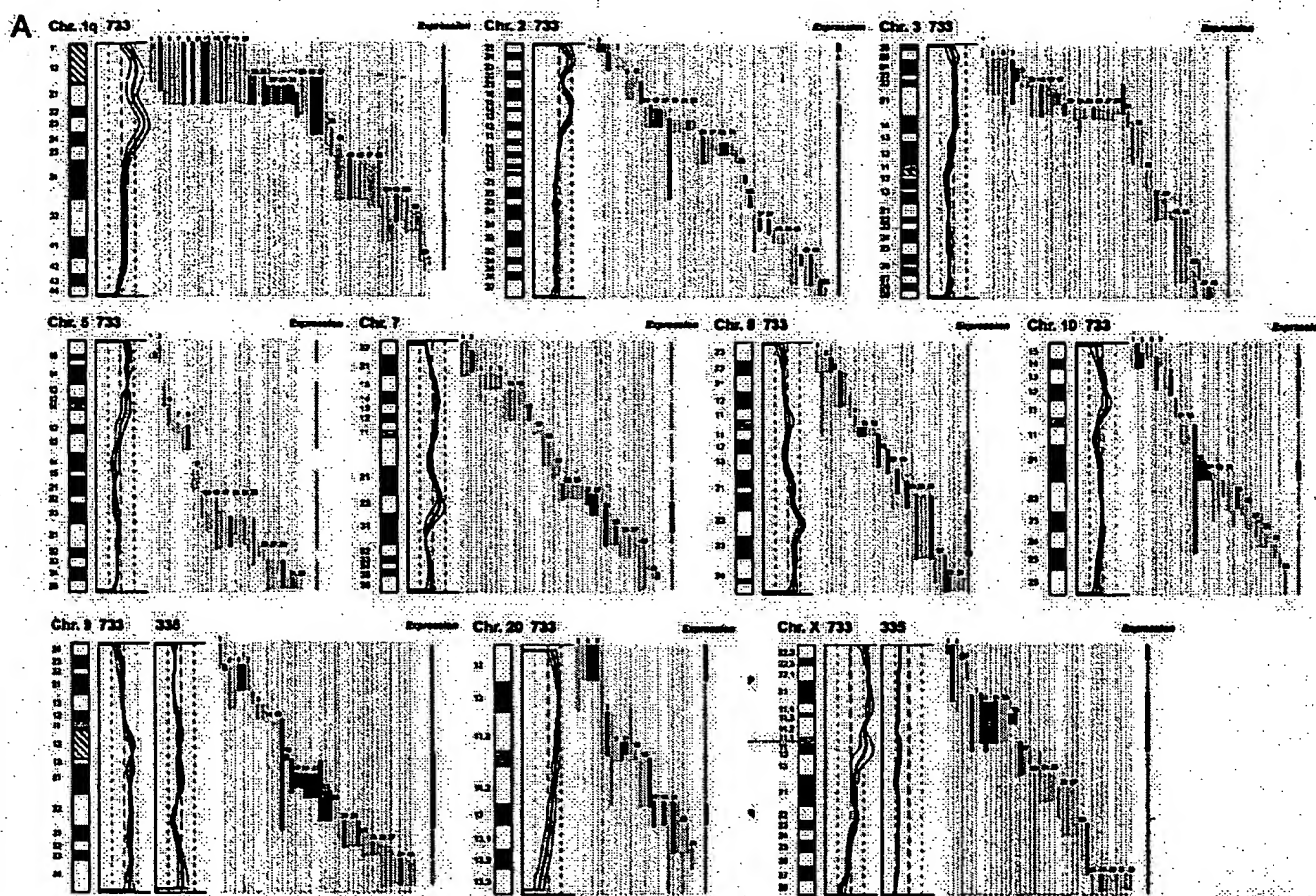
**Material**—Bladder tumor biopsies were sampled after informed consent was obtained and after removal of tissue for routine pathology examination. By light microscopy tumors 335 and 532 were staged by an experienced pathologist as pTa (superficial papillary),

From the ‡Department of Clinical Biochemistry, Molecular Diagnostic Laboratory and \*\*Department of Urology, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark, §AROS Applied Biotechnology ApS, Gustav Wiedesvej 10, DK-8000 Aarhus C, Denmark, ||UCSF Cancer Center and Department of Laboratory Medicine, University of California, San Francisco, CA 94143-0808, and ‡‡Institute of Medical Biochemistry and Danish Centre for Human Genome Research, Ole Worms Allé 170, Aarhus University, DK-8000 Aarhus C, Denmark

Received, September 26, 2001, and in revised form, November 7, 2001

Published, MCP Papers in Press, November 13, 2001, DOI 10.1074/mcp.M100019-MCP200

<sup>1</sup> The abbreviations used are: CGH, comparative genomic hybridization; TCC, transitional cell carcinoma; LOH, loss of heterozygosity; PA-FABP, psoriasis-associated fatty acid-binding protein; 2D, two-dimensional.



**Fig. 1. DNA copy number and mRNA expression level.** Shown from left to right are chromosome (Chr.), CGH profiles, gene location and expression level of specific genes, and overall expression level along the chromosome. **A**, expression of mRNA in invasive tumor 733 as compared with the non-invasive counterpart tumor 335. **B**, expression of mRNA in invasive tumor 827 compared with the non-invasive counterpart tumor 532. The average fluorescent signal ratio between tumor DNA and normal DNA is shown along the length of the chromosome (left). The bold curve in the ratio profile represents a mean of four chromosomes and is surrounded by thin curves indicating one standard deviation. The central vertical line (broken) indicates a ratio value of 1 (no change), and the vertical lines next to it (dotted) indicate a ratio of 0.5 (left) and 2.0 (right). In chromosomes where the non-invasive tumor 335 used for comparison showed alterations in DNA content, the ratio profile of that chromosome is shown to the right of the invasive tumor profile. The colored bars represent one gene each, identified by the running numbers above the bars (the name of the gene can be seen at [www.MDL.DK/sdata.html](http://www.MDL.DK/sdata.html)). The bars indicate the purported location of the gene, and the colors indicate the expression level of the gene in the invasive tumor compared with the non-invasive counterpart; >2-fold increase (black), >2-fold decrease (blue), no significant change (orange). The bar to the far right, entitled *Expression* shows the resulting change in expression along the chromosome; the colors indicate that at least half of the genes were up-regulated (black), at least half of the genes down-regulated (blue), or more than half of the genes are unchanged (orange). If a gene was absent in one of the samples and present in another, it was regarded as more than a 2-fold change. A 2-fold level was chosen as this corresponded to one standard deviation in a double determination of ~1800 genes. Centromeres and heterochromatic regions were excluded from data analysis.

grade I and II, respectively, tumors 733 and 827 were staged as pT1 (invasive into submucosa), 733 was staged as solid, and 827 was staged as papillary, both grade III.

**mRNA Preparation**—Tissue biopsies, obtained fresh from surgery, were embedded immediately in a sodium-guanidinium thiocyanate solution and stored at  $-80^{\circ}\text{C}$ . Total RNA was isolated using the RNeasy B RNA isolation method (WAK-Chemie Medical GMBH). poly(A)<sup>+</sup> RNA was isolated by an oligo(dT) selection step (Oligotex mRNA kit; Qiagen).

**cRNA Preparation**—1  $\mu\text{g}$  of mRNA was used as starting material. The first and second strand cDNA synthesis was performed using the SuperScript<sup>®</sup> choice system (Invitrogen) according to the manufacturer's instructions but using an oligo(dT) primer containing a T7 RNA polymerase binding site. Labeled cRNA was prepared using the ME-GAScript<sup>®</sup> *in vitro* transcription kit (Ambion). Biotin-labeled CTP and

UTP (Enzo) was used, together with unlabeled NTPs in the reaction. Following the *in vitro* transcription reaction, the unincorporated nucleotides were removed using RNeasy columns (Qiagen).

**Array Hybridization and Scanning**—Array hybridization and scanning was modified from a previous method (13). 10  $\mu\text{g}$  of cRNA was fragmented at  $94^{\circ}\text{C}$  for 35 min in buffer containing 40 mM Tris acetate, pH 8.1, 100 mM KOAc, 30 mM MgOAc. Prior to hybridization, the fragmented cRNA in a 6 $\times$  SSPE-T hybridization buffer (1 M NaCl, 10 mM Tris, pH 7.6, 0.005% Triton), was heated to  $95^{\circ}\text{C}$  for 5 min, subsequently cooled to  $40^{\circ}\text{C}$ , and loaded onto the Affymetrix probe array cartridge. The probe array was then incubated for 16 h at  $40^{\circ}\text{C}$  at constant rotation (60 rpm). The probe array was exposed to 10 washes in 6 $\times$  SSPE-T at  $25^{\circ}\text{C}$  followed by 4 washes in 0.5 $\times$  SSPE-T at  $50^{\circ}\text{C}$ . The biotinylated cRNA was stained with a streptavidin-phycoerythrin conjugate, 10  $\mu\text{g}/\text{ml}$  (Molecular Probes) in 6 $\times$  SSPE-T



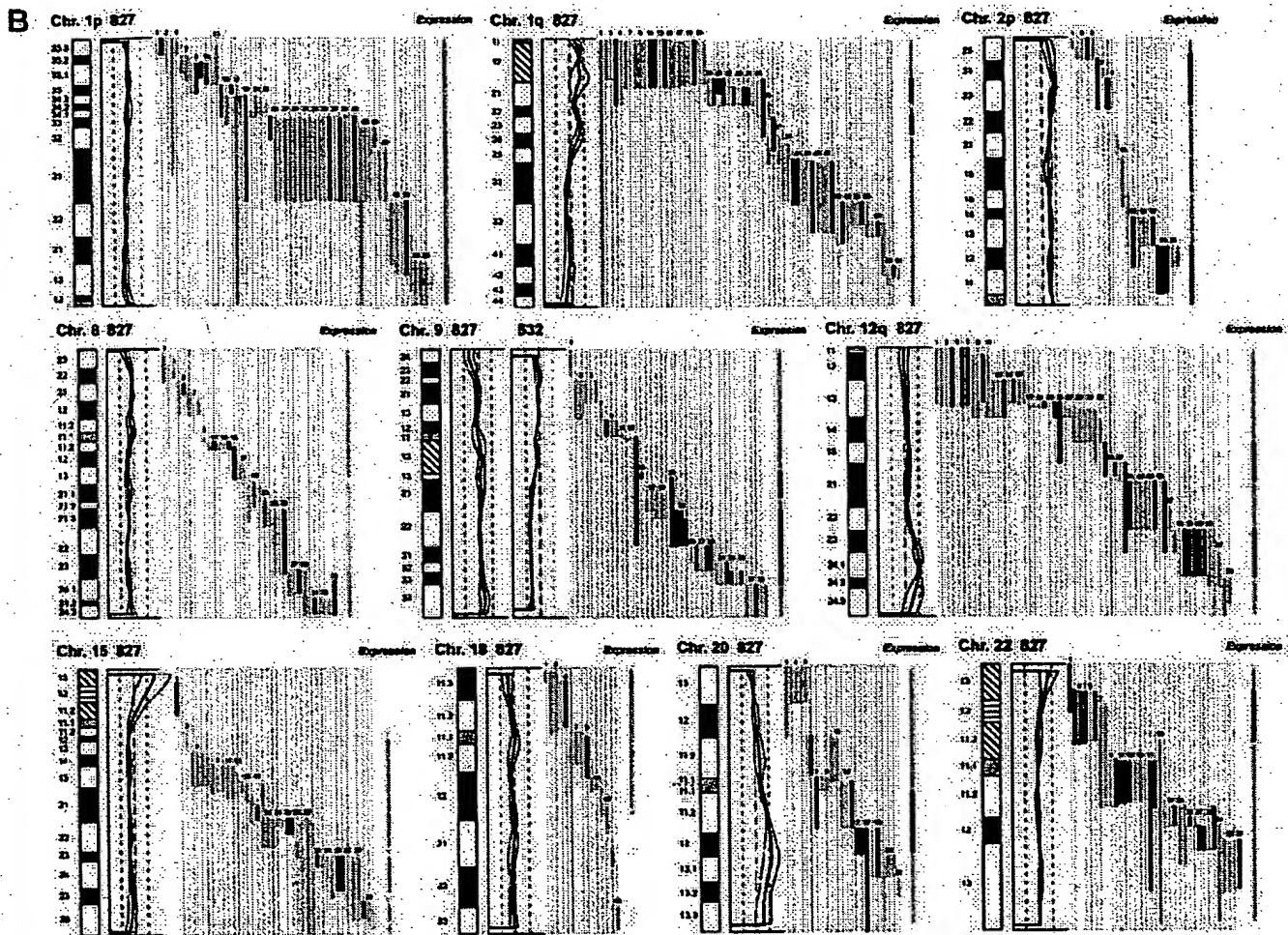


FIG. 1—continued

for 30 min at 25 °C followed by 10 washes in 6× SSPE-T at 25 °C. The probe arrays were scanned at 560 nm using a confocal laser scanning microscope (made for Affymetrix by Hewlett-Packard). The readings from the quantitative scanning were analyzed by Affymetrix gene expression analysis software.

**Microsatellite Analysis**—Microsatellite Analysis was performed as described previously (14). Microsatellites were selected by use of [www.ncbi.nlm.nih.gov/genemap98](http://www.ncbi.nlm.nih.gov/genemap98), and primer sequences were obtained from the genome data base at [www.gdb.org](http://www.gdb.org). DNA was extracted from tumor and blood and amplified by PCR in a volume of 20  $\mu$ l for 35 cycles. The amplicons were denatured and electrophoresed for 3 h in an ABI Prism 377. Data were collected in the Gene Scan program for fragment analysis. Loss of heterozygosity was defined as less than 33% of one allele detected in tumor amplicons compared with blood.

**Proteomic Analysis**—TCCs were minced into small pieces and homogenized in a small glass homogenizer in 0.5 ml of lysis solution. Samples were stored at -20 °C until use. The procedure for 2D gel electrophoresis has been described in detail elsewhere (15, 16). Gels were stained with silver nitrate and/or Coomassie Brilliant Blue. Proteins were identified by a combination of procedures that included microsequencing, mass spectrometry, two-dimensional gel Western immunoblotting, and comparison with the master two-dimensional gel image of human keratinocyte proteins; see [biobase.dk/cgi-bin/celis](http://biobase.dk/cgi-bin/celis).

**CGH**—Hybridization of differentially labeled tumor and normal DNA to normal metaphase chromosomes was performed as described previously (10). Fluorescein-labeled tumor DNA (200 ng), Texas Red-

labeled reference DNA (200 ng); and human Cot-1 DNA (20  $\mu$ g) were denatured at 37 °C for 5 min and applied to denatured normal metaphase slides. Hybridization was at 37 °C for 2 days. After washing, the slides were counterstained with 0.15  $\mu$ g/ml 4,6-diamidino-2-phenylindole in an anti-fade solution. A second hybridization was performed for all tumor samples using fluorescein-labeled reference DNA and Texas Red-labeled tumor DNA (inverse labeling) to confirm the aberrations detected during the initial hybridization. Each CGH experiment also included a normal control hybridization using fluorescein- and Texas Red-labeled normal DNA. Digital image analysis was used to identify chromosomal regions with abnormal fluorescence ratios, indicating regions of DNA gains and losses. The average green:red fluorescence intensity ratio profiles were calculated using four images of each chromosome (eight chromosomes total) with normalization of the green:red fluorescence intensity ratio for the entire metaphase and background correction. Chromosome identification was performed based on 4,6-diamidino-2-phenylindole banding patterns. Only images showing uniform high intensity fluorescence with minimal background staining were analyzed. All centromeres, p arms of acrocentric chromosomes, and heterochromatic regions were excluded from the analysis.

## RESULTS

**Comparative Genomic Hybridization**—The CGH analysis identified a number of chromosomal gains and losses in the

TABLE I  
Correlation between alterations detected by CGH and by expression monitoring

Top, CGH used as independent variable (if CGH alteration – what expression ratio was found); bottom, altered expression used as independent variable (if expression alteration – what CGH deviation was found).

CGH alterations	Tumor 733 vs. 335		CGH alterations	Tumor 827 vs. 532	
	Expression change clusters	Concordance		Expression change clusters	Concordance
13 Gain	10 Up-regulation 0 Down-regulation 3 No change	77%	10 Gain	8 Up-regulation 0 Down-regulation 2 No change	80%
10 Loss	1 Up-regulation 5 Down-regulation 4 No change	50%	12 Loss	3 Up-regulation 2 Down regulation 7 No change	17%
Expression change clusters	Tumor 733 vs. 335		Expression change clusters	Tumor 827 vs. 532	
	CGH alterations	Concordance		CGH alterations	Concordance
16 Up-regulation	11 Gain 2 Loss 3 No change	69%	17 Up-regulation	10 Gain 5 Loss 2 No change	59%
21 Down-regulation	1 Gain 8 Loss 12 No change	38%	9 Down-regulation	0 Gain 3 Loss 6 No change	33%
15 No change	3 Gain 3 Loss 9 No change	60%	21 No change	1 Gain 3 Loss 17 No change	81%

two invasive tumors (stage pT1, TCCs 733 and 827), whereas the two non-invasive papillomas (stage pTa, TCCs 335 and 532) showed only 9p–, 9q22–q33–, and X–, and 7+, 9q–, and Y–, respectively. Both invasive tumors showed changes (1q22–24+, 2q14.1–qter–, 3q12–q13.3–, 6q12–q22–, 9q34+, 11q12–q13+, 17+, and 20q11.2–q12+) that are typical for their disease stage, as well as additional alterations, some of which are shown in Fig. 1. Areas with gains and losses deviated from the normal copy number to some extent, and the average numerical deviation from normal was 0.4-fold in the case of TCC 733 and 0.3-fold for TCC 827. The largest changes, amounting to at least a doubling of chromosomal content, were observed at 1q23 in TCC 733 (Fig. 1A) and 20q12 in TCC 827 (Fig. 1B).

**mRNA Expression in Relation to DNA Copy Number**—The mRNA levels from the two invasive tumors (TCCs 827 and 733) were compared with the two non-invasive counterparts (TCCs 532 and 335). This was done in two separate experiments in which we compared TCCs 733 to 335 and 827 to 532, respectively, using two different scaling settings for the arrays to rule out scaling as a confounding parameter. Approximately 1,800 genes that yielded a signal on the arrays were searched in the Unigene and Genemap data bases for chromosomal location, and those with a known location (1096) were plotted as bars covering their purported locus. In that way it was possible to construct a graphic presentation of DNA copy number and relative mRNA levels along the individual chromosomes (Fig. 1).

For each mRNA a ratio was calculated between the level in the invasive versus the non-invasive counterpart. Bars, which represent chromosomal location of a gene, were color-coded according to the expression ratio, and only differences larger

than 2-fold were regarded as informative (Fig. 1). The density of genes along the chromosomes varied, and areas containing only one gene were excluded from the calculations. The resolution of the CGH method is very low, and some of the outlier data may be because of the fact that the boundaries of the chromosomal aberrations are not known at high resolution.

Two sets of calculations were made from the data. For the first set we used CGH alterations as the independent variable and estimated the frequency of expression alterations in these chromosomal areas. In general, areas with a strong gain of chromosomal material contained a cluster of genes having increased mRNA expression. For example, both chromosomes 1q21–q25, 2p and 9q, showed a relative gain of more than 100% in DNA copy number that was accompanied by increased mRNA expression levels in the two tumor pairs (Fig. 1). In most cases, chromosomal gains detected by CGH were accompanied by an increased level of transcripts in both TCCs 733 (77%) and 827 (80%) (Table I, top). Chromosomal losses, on the other hand, were not accompanied by decreased expression in several cases, and were often registered as having unaltered RNA levels (Table I, top). The inability to detect RNA expression changes in these cases was not because of fewer genes mapping to the lost regions (data not shown).

In the second set of calculations we selected expression alterations above 2-fold as the independent variable and estimated the frequency of CGH alterations in these areas. As above, we found that increased transcript expression correlated with gain of chromosomal material (TCC 733, 69% and TCC 827, 59%), whereas reduced expression was often detected in areas with unaltered CGH ratios (Table I, bottom). Furthermore, as a control we looked at areas with no alter-

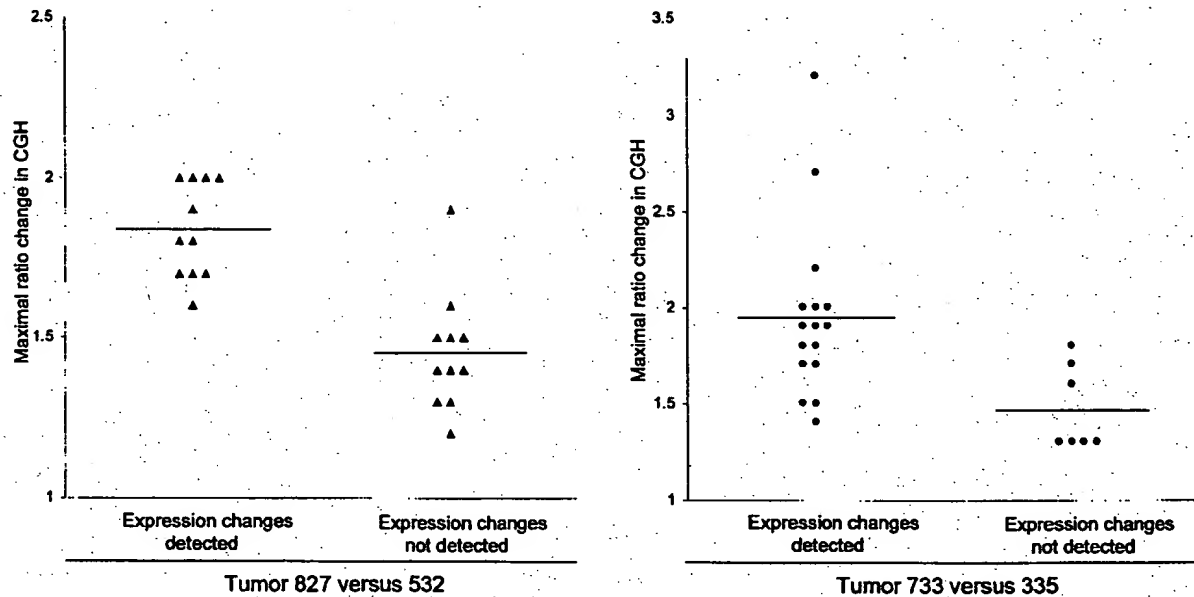


Fig. 2. Correlation between maximum CGH aberration and the ability to detect expression change by oligonucleotide array monitoring. The aberration is shown as a numerical -fold change in ratio between invasive tumors 827 (▲) and 733 (◆) and their non-invasive counterparts 532 and 335. The expression change was taken from the *Expression* line to the *right* in Fig. 1, which depicts the resulting expression change for a given chromosomal region. At least half of the mRNAs from a given region have to be either up- or down-regulated to be scored as an expression change. All chromosomal arms in which the CGH ratio plus or minus one standard deviation was outside the ratio value of one were included.

ation in expression. No alteration was detected by CGH in most of these areas (TCC 733, 60% and TCC 827, 81%; see Table I, *bottom*). Because the ability to observe reduced or increased mRNA expression clustering to a certain chromosomal area clearly reflected the extent of copy number changes, we plotted the maximum CGH aberrations in the regions showing CGH changes against the ability to detect a change in mRNA expression as monitored by the oligonucleotide arrays (Fig. 2). For both tumors TCC 733 ( $p < 0.015$ ) and TCC 827 ( $p < 0.00003$ ) a highly significant correlation was observed between the level of CGH ratio change (reflecting the DNA copy number) and alterations detected by the array based technology (Fig. 2). Similar data were obtained when areas with altered expression were used as independent variables. These areas correlated best with CGH when the CGH ratio deviated 1.6- to 2.0-fold (Table I, *bottom*) but mostly did not at lower CGH deviations. These data probably reflect that loss of an allele may only lead to a 50% reduction in expression level, which is at the cut-off point for detection of expression alterations. Gain of chromosomal material can occur to a much larger extent.

**Microsatellite-based Detection of Minor Areas of Losses**—In TCC 733, several chromosomal areas exhibiting DNA amplification were preceded or followed by areas with a normal CGH but reduced mRNA expression (see Fig. 1, TCC 733 chromosome 1q32, 2p21, and 7q21 and q32, 9q34, and 10q22). To determine whether these results were because of undetected loss of chromosomal material in these regions or

because of other non-structural mechanisms regulating transcription, we examined two microsatellites positioned at chromosome 1q25–32 and two at chromosome 2p22. Loss of heterozygosity (LOH) was found at both 1q25 and at 2p22 indicating that minor deleted areas were not detected with the resolution of CGH (Fig. 3). Additionally, chromosome 2p in TCC 733 showed a CGH pattern of gain/no change/gain of DNA that correlated with transcript increase/decrease/increase. Thus, for the areas showing increased expression there was a correlation with the DNA copy number alterations (Fig. 1A). As indicated above, the mRNA decrease observed in the middle of the chromosomal gain was because of LOH, implying that one of the mechanisms for mRNA down-regulation may be regions that have undergone smaller losses of chromosomal material. However, this cannot be detected with the resolution of the CGH method.

In both TCC 733 and TCC 827, the telomeric end of chromosome 11p showed a normal ratio in the CGH analysis; however, clusters of five and three genes, respectively, lost their expression. Two microsatellites (D11S1760, D11S922) positioned close to MUC2, IGF2, and cathepsin D indicated LOH as the most likely mechanism behind the loss of expression (data not shown).

A reduced expression of mRNA observed in TCC 733 at chromosomes 3q24, 11p11, 12p12.2, 12q21.1, and 16q24 and in TCC 827 at chromosome 11p15.5, 12p11, 15q11.2, and 18q12 was also examined for chromosomal losses using microsatellites positioned as close as possible to the gene loci



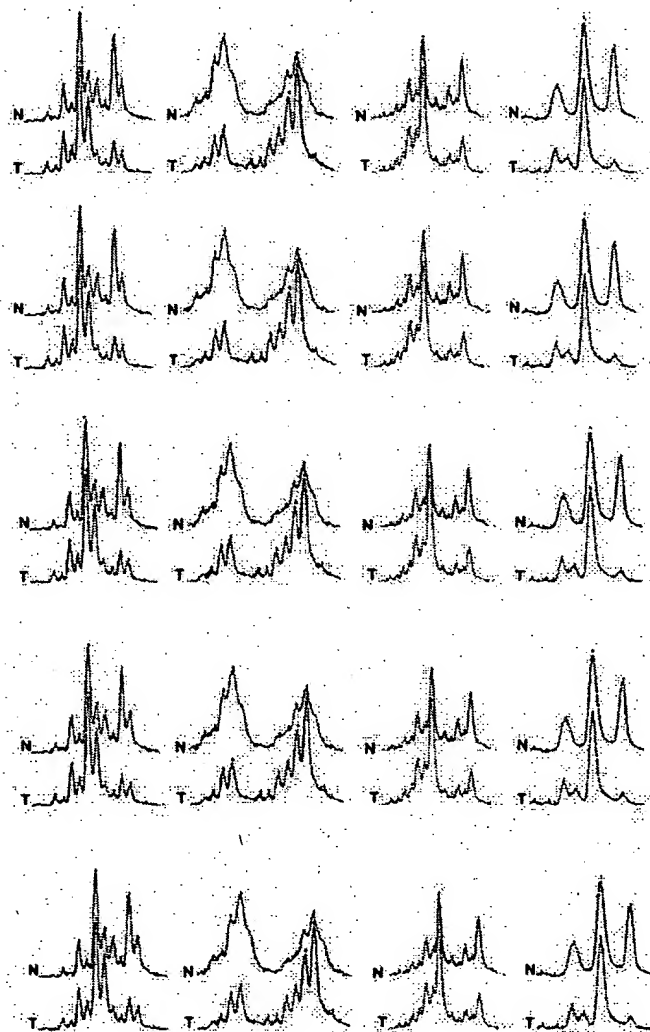


FIG. 3. Microsatellite analysis of loss of heterozygosity. Tumor 733 showing loss of heterozygosity at chromosome 1q25, detected (a) by D1S215 close to HLA class I histocompatibility antigen (gene number 38 in Fig. 1), (b) by D1S2735 close to cathepsin E (gene number 41 in Fig. 1), and (c) at chromosome 2p23 by D2S2251 close to general  $\beta$ -spectrin (gene number 11 on Fig. 1) and of (d) tumor 827 showing loss of heterozygosity at chromosome 18q12 by S18S1118 close to mitochondrial 3-oxoacyl-coenzyme A thiolase (gene number 12 in Fig. 1). The upper curves show the electropherogram obtained from normal DNA from leukocytes (N), and the lower curves show the electropherogram from tumor DNA (T). In all cases one allele is partially lost in the tumor amplicon.

showing reduced mRNA transcripts. Only the microsatellite positioned at 18q12 showed LOH (Fig. 3), suggesting that transcriptional down-regulation of genes in the other regions may be controlled by other mechanisms.

**Relation between Changes in mRNA and Protein Levels—**2D-PAGE analysis, in combination with Coomassie Brilliant Blue and/or silver staining, was carried out on all four tumors using fresh biopsy material. 40 well resolved abundant known proteins migrating in areas away from the edges of the pH

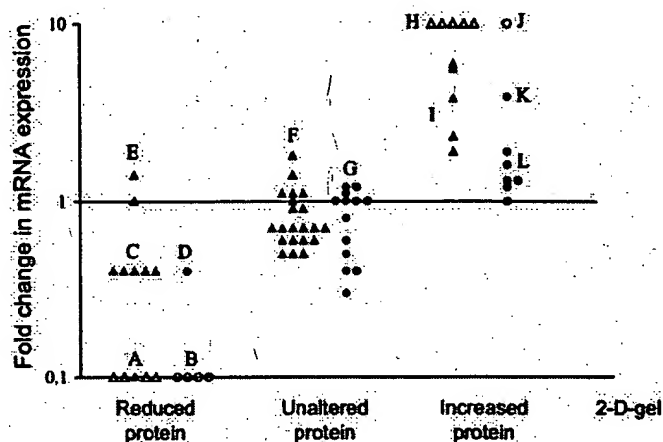
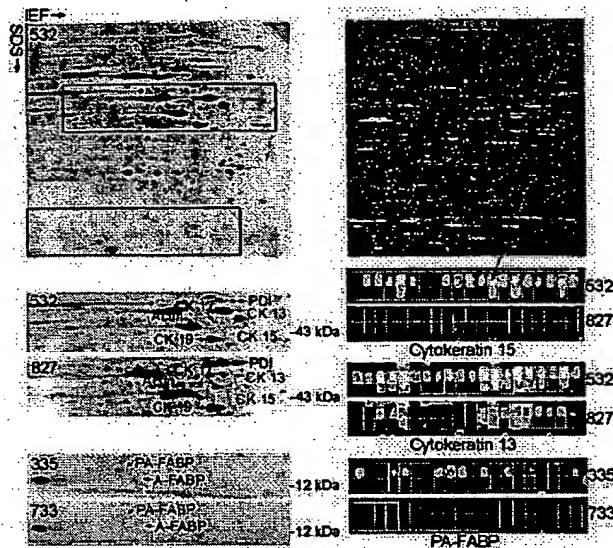


FIG. 4. Correlation between protein levels as judged by 2D-PAGE and transcript ratio. For comparison proteins were divided in three groups, unaltered in level or up- or down-regulated (horizontal axis). The mRNA ratio as determined by oligonucleotide arrays was plotted for each gene (vertical axis). ▲, mRNAs that were scored as present in both tumors used for the ratio calculation; ●, mRNAs that were scored as absent in the invasive tumors (along horizontal axis) or as absent in non-invasive reference (top of figure). Two different scalings were used to exclude scaling as a confounder, TCCs 827 and 532 (▲▲) were scaled with background suppression, and TCCs 733 and 335 (●●) were scaled without suppression. Both comparisons showed highly significant ( $p < 0.005$ ) differences in mRNA ratios between the groups. Proteins shown were as follows: Group A (from left), phosphoglucomutase 1, glutathione transferase class  $\mu$  number 4, fatty acid-binding protein homologue, cytokeratin 15, and cytokeratin 13; B (from left), fatty acid-binding protein homologue, 28-kDa heat shock protein, cytokeratin 13, and calcyclin; C (from left),  $\alpha$ -enolase, hnRNP B1, 28-kDa heat shock protein, 14-3-3- $\epsilon$ , and pre-mRNA splicing factor; D, mesothelial keratin K7 (type II); E (from top), glutathione S-transferase- $\pi$  and mesothelial keratin K7 (type II); F (from top and left), adenyl cyclase-associated protein, E-cadherin, keratin 19, calgizzarin, phosphoglycerate mutase, annexin IV, cytoskeletal  $\gamma$ -actin, hnRNP A1, integral membrane protein calnexin (IP90), hnRNP H, brain-type clathrin light chain-a, hnRNP F, 70-kDa heat shock protein, heterogeneous nuclear ribonucleoprotein A/B, translationally controlled tumor protein, liver glyceraldehyde-3-phosphate dehydrogenase, keratin 8, aldehyde reductase, and Na,K-ATPase  $\beta$ -1 subunit; G, (from top and left), TCP20, calgizzarin, 70-kDa heat shock protein, calnexin, hnRNP H, cytokeratin 15, ATP synthase, keratin 19, triosephosphate isomerase, hnRNP F, liver glyceraldehyde-3-phosphatase dehydrogenase, glutathione S-transferase- $\pi$ , and keratin 8; H (from left), plasma gelsolin, autoantigen calreticulin, thioredoxin, and NAD $^{+}$ -dependent 15 hydroxyprostaglandin dehydrogenase; I (from top), prollyl 4-hydroxylase  $\beta$ -subunit, cytokeratin 20, cytokeratin 17, prothibition, and fructose 1,6-biphosphatase; J annexin II; K, annexin IV; L (from top and left), 90-kDa heat shock protein, prollyl 4-hydroxylase  $\beta$ -subunit,  $\alpha$ -enolase, GRP 78, cyclophilin, and cofilin.

gradient, and having a known chromosomal location, were selected for analysis in the TCC pair 827/532. Proteins were identified by a combination of methods (see "Experimental Procedures"). In general there was a highly significant correlation ( $p < 0.005$ ) between mRNA and protein alterations (Fig. 4). Only one gene showed disagreement between transcript alteration and protein alteration. Except for a group of cyto-



**FIG. 5. Comparison of protein and transcript levels in invasive and non-invasive TCCs.** The upper part of the figure shows a 2D gel (left) and the oligonucleotide array (right) of TCC 532. The red rectangles on the upper gel highlight the areas that are compared below. Identical areas of 2D gels of TCCs 532 and 827 are shown below. Clearly, cytokeratins 13 and 15 are strongly down-regulated in TCC 827 (red annotation). The tile on the array containing probes for cytokeratin 15 is enlarged below the array (red arrow) from TCC 532 and is compared with TCC 827. The upper row of squares in each tile corresponds to perfect match probes; the lower row corresponds to mismatch probes containing a mutation (used for correction for unspecific binding). Absence of signal is depicted as black, and the higher the signal the lighter the color. A high transcript level was detected in TCC 532 (6151 units) whereas a much lower level was detected in TCC 827 (absence of signals). For cytokeratin 13, a high transcript level was also present in TCC 532 (15659 units), and a much lower level was present in TCC 827 (623 units). The 2D gels at the bottom of the figure (left) show levels of PA-FABP and adipocyte-FABP in TCCs 335 and 733 (invasive), respectively. Both proteins are down-regulated in the invasive tumor. To the right we show the array tiles for the PA-FABP transcript. A medium transcript level was detected in the case of TCC 335 (1277 units) whereas very low levels were detected in TCC 733 (166 units). IEF, isoelectric focusing.

keratins encoded by genes on chromosome 17 (Fig. 5) the analyzed proteins did not belong to a particular family. 26 well focused proteins whose genes had a known chromosomal location were detected in TCCs 733 and 335, and of these 19 correlated ( $p < 0.005$ ) with the mRNA changes detected using the arrays (Fig. 4). For example, PA-FABP was highly expressed in the non-invasive TCC 335 but lost in the invasive counterpart (TCC 733; see Fig. 5). The smaller number of proteins detected in both 733 and 335 was because of the smaller size of the biopsies that were available.

11 chromosomal regions where CGH showed aberrations that corresponded to the changes in transcript levels also showed corresponding changes in the protein level (Table II). These regions included genes that encode proteins that are found to be frequently altered in bladder cancer, namely cytokeratins 17 and 20, annexins II and IV, and the fatty acid-binding proteins PA-FABP and FBP1. Four of these proteins were encoded by genes in chromosome 17q, a frequently amplified chromosomal area in invasive bladder cancers.

#### DISCUSSION

Most human cancers have abnormal DNA content, having lost some chromosomal parts and gained others. The present study provides some evidence as to the effect of these gains and losses on gene expression in two pairs of non-invasive and invasive TCCs using high throughput expression arrays and proteomics, in combination with CGH. In general, the results showed that there is a clear individual regulation of the mRNA expression of single genes, which in some cases was superimposed by a DNA copy number effect. In most cases, genes located in chromosomal areas with gains often exhibited increased mRNA expression, whereas areas showing losses showed either no change or a reduced mRNA expression. The latter might be because of the fact that losses most often are restricted to loss of one allele, and the cut-off point for detection of expression alterations was a 2-fold change, thus being at the border of detection. In several cases, how-

**TABLE II**  
Proteins whose expression level correlates with both mRNA and gene dose changes

Protein	Chromosomal location	Tumor TCC	CGH alteration	Transcript alteration <sup>a</sup>	Protein alteration
Annexin II	1q21	733	Gain	Abs to Pres <sup>a</sup>	Increase
Annexin IV	2p13	733	Gain	3.9-Fold up	Increase
Cytokeratin 17	17q12-q21	827	Gain	3.8-Fold up	Increase
Cytokeratin 20	17q21.1	827	Gain	5.6-Fold up	Increase
(PA-)FABP	8q21.2	827	Loss	10-Fold down	Decrease
FBP1	9q22	827	Gain	2.3-Fold up	Increase
Plasma gelsolin	9q31	827	Gain	Abs to Pres	Increase
Heat shock protein 28	15q12-q13	827	Loss	2.5-Fold up	Decrease
Prohibitin	17q21	827/733	Gain	3.7-/2.5-Fold up <sup>b</sup>	Increase
Prolyl-4-hydroxyl	17q25	827/733	Gain	5.7-/1.6-Fold up	Increase
hnRNPB1	7p15	827	Loss	2.5-Fold down	Decrease

<sup>a</sup> Abs, absent; Pres, present.

<sup>b</sup> In cases where the corresponding alterations were found in both TCCs 827 and 733 these are shown as 827/733.

ever, an increase or decrease in DNA copy number was associated with *de novo* occurrence or complete loss of transcript, respectively. Some of these transcripts could not be detected in the non-invasive tumor but were present at relatively high levels in areas with DNA amplifications in the invasive tumors (e.g. in TCC 733 transcript from cellular ligand of annexin II gene (chromosome 1q21) from absent to 2670 arbitrary units; in TCC 827 transcript from small proline-rich protein 1 gene (chromosome 1q12-q21.1) from absent to 1326 arbitrary units). It may be anticipated from these data that significant clustering of genes with an increased expression to a certain chromosomal area indicates an increased likelihood of gain of chromosomal material in this area.

Considering the many possible regulatory mechanisms acting at the level of transcription, it seems striking that the gene dose effects were so clearly detectable in gained areas. One hypothetical explanation may lie in the loss of controlled methylation in tumor cells (17–19). Thus, it may be possible that in chromosomes with increased DNA copy numbers two or more alleles could be demethylated simultaneously leading to a higher transcription level, whereas in chromosomes with losses the remaining allele could be partly methylated, turning off the process (20, 21). A recent report has documented a ploidy regulation of gene expression in yeast, but in this case all the genes were present in the same ratio (22), a situation that is not analogous to that of cancer cells, which show marked chromosomal aberrations, as well as gene dosage effects.

Several CGH studies of bladder cancer have shown that some chromosomal aberrations are common at certain stages of disease progression, often occurring in more than 1 of 3 tumors. In pTa tumors, these include 9p–, 9q–, 1q+, Y– (2, 6), and in pT1 tumors, 2q–, 11p–, 11q–, 1q+, 5p+, 8q+, 17q+, and 20q+ (2–4, 6, 7). The pTa tumors studied here showed similar aberrations such as 9p– and 9q22–q33– and 9q– and Y–, respectively. Likewise, the two minimal invasive pT1 tumors showed aberrations that are commonly seen at that stage, and TCC 827 had a remarkable resemblance to the commonly seen pattern of losses and gains, such as 1q22–24 amplification (seen in both tumors), 11q14–q22 loss, the latter often linked to 17q+ (both tumors), and 1q+ and 9p–, often linked to 20q+ and 11q13+ (both tumors) (7–9). These observations indicate that the pairs of tumors used in this study exhibit chromosomal changes observed in many tumors, and therefore the findings could be of general importance for bladder cancer.

Considering that the mapping resolution of CGH is of about 20 megabases it is only possible to get a crude picture of chromosomal instability using this technique. Occasionally, we observed reduced transcript levels close to or inside regions with increased copy numbers. Analysis of these regions by positioning heterozygous microsatellites as close as possible to the locus showing reduced gene expression revealed loss of heterozygosity in several cases. It seems likely that multiple and different events occur along each chromosomal

arm and that the use of cDNA microarrays for analysis of DNA copy number changes will reach a resolution that can resolve these changes, as has recently been proposed (2). The outlier data were not more frequent at the boundaries of the CGH aberrations. At present we do not know the mechanism behind chromosomal aneuploidy and cannot predict whether chromosomal gains will be transcribed to a larger extent than the two native alleles. A mechanism as genetic imprinting has an impact on the expression level in normal cells and is often reduced in tumors. However, the relation between imprinting and gain of chromosomal material is not known.

We regard it as a strength of this investigation that we were able to compare invasive tumors to benign tumors rather than to normal urothelium, as the tumors studied were biologically very close and probably may represent successive steps in the progression of bladder cancer. Despite the limited amount of fresh tissue available it was possible to apply three different state of the art methods. The observed correlation between DNA copy number and mRNA expression is remarkable when one considers that different pieces of the tumor biopsies were used for the different sets of experiments. This indicates that bladder tumors are relatively homogenous, a notion recently supported by CGH and LOH data that showed a remarkable similarity even between tumors and distant metastasis (10, 23).

In the few cases analyzed, mRNA and protein levels showed a striking correspondence although in some cases we found discrepancies that may be attributed to translational regulation, post-translational processing, protein degradation, or a combination of these. Some transcripts belong to undertranslated mRNA pools, which are associated with few translationally inactive ribosomes; these pools, however, seem to be rare (24). Protein degradation, for example, may be very important in the case of polypeptides with a short half-life (e.g. signaling proteins). A poor correlation between mRNA and protein levels was found in liver cells as determined by arrays and 2D-PAGE (25), and a moderate correlation was recently reported by Ideker *et al.* (26) in yeast.

Interestingly, our study revealed a much better correlation between gained chromosomal areas and increased mRNA levels than between loss of chromosomal areas and reduced mRNA levels. In general, the level of CGH change determined the ability to detect a change in transcript. One possible explanation could be that by losing one allele the change in mRNA level is not so dramatic as compared with gain of material, which can be rather unlimited and may lead to a severalfold increase in gene copy number resulting in a much higher impact on transcript level. The latter would be much easier to detect on the expression arrays as the cut-off point was placed at a 2-fold level so as not to be biased by noise on the array. Construction of arrays with a better signal to noise ratio may in the future allow detection of lesser than 2-fold alterations in transcript levels, a feature that may facilitate the analysis of the effect of loss of chromosomal areas on transcript levels.

In eleven cases we found a significant correlation between DNA copy number, mRNA expression, and protein level. Four of these proteins were encoded by genes located at a frequently amplified area in chromosome 17q. Whether DNA copy number is one of the mechanisms behind alteration of these eleven proteins is at present unknown and will have to be proved by other methods using a larger number of samples. One factor making such studies complicated is the large extent of protein modification that occurs after translation, requiring immunoidentification and/or mass spectrometry to correctly identify the proteins in the gels.

In conclusion, the results presented in this study exemplify the large body of knowledge that may be possible to gather in the future by combining state of the art techniques that follow the pathway from DNA to protein (26). Here, we used a traditional chromosomal CGH method, but in the future high resolution CGH based on microarrays with many thousand radiation hybrid-mapped genes will increase the resolution and information derived from these types of experiments (2). Combined with expression arrays analyzing transcripts derived from genes with known locations, and 2D gel analysis to obtain information at the post-translational level, a clearer and more developed understanding of the tumor genome will be forthcoming.

**Acknowledgments**—We thank Mie Madsen, Hanne Steen, Inge Lis Thorsen, Hans Lund, Vikolaj Ørntoft, and Lynn Bjerke for technical help and Thomas Gingeras, Christine Harrington, and Morten Østergaard for valuable discussions.

\* This work was supported by grants from The Danish Cancer Society, the University of Aarhus, Aarhus County, Novo Nordic, the Danish Biotechnology Program, the Frenkels Foundation, the John and Birthe Meyer Foundation, and NCI, National Institutes of Health Grant CA47537. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Dept. of Clinical Biochemistry, Molecular Diagnostic Laboratory, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark. Tel.: 45-89495100/45-86156201 (private); Fax: 45-89496018; E-mail: orntoft@kba.sks.au.dk.

## REFERENCES

- Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature* 396, 643–649.
- Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* 23, 41–46.
- de Cremoux, P., Martin, E. C., Vincent-Salomon, A., Dieras, V., Barbaroux, C., Liva, S., Pouillart, P., Sastre-Garau, X., and Magdelenat, H. (1999) Quantitative PCR analysis of c-erb B-2 (HER2/neu) gene amplification and comparison with p185(HER2/neu) protein expression in breast cancer drill biopsies. *Int. J. Cancer* 83, 157–161.
- Brungier, P. P., Tamimi, Y., Shuuring, E., and Schalken, J. (1996) Expression of cyclin D1 and EMS1 in bladder tumors; relationship with chromosome 11q13 amplifications. *Oncogene* 12, 1747–1753.
- Slavc, I., Ellenbogen, R., Jung, W. H., Vawter, G. F., Kretschmar, C., Grier, H., and Korf, B. R. (1990) *myc* gene amplification and expression in primary human neuroblastoma. *Cancer Res.* 50, 1459–1463.
- Sauter, G., Carroll, P., Moch, H., Kallioniemi, A., Kerschmann, R., Narayan, P., Mihatsch, M. J., and Waldman, F. M. (1995) *c-myc* copy number gains in bladder cancer detected by fluorescence *in situ* hybridization. *Am. J. Pathol.* 146, 1131–1139.
- Richter, J., Jiang, F., Gorog, J. P., Sartorius, G., Egenter, C., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1997) Marked genetic differences between stage pTa and stage pT1 papillary bladder cancer detected by comparative genomic hybridization. *Cancer Res.* 57, 2860–2864.
- Richter, J., Beffa, L., Wagner, U., Schraml, P., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1998) Patterns of chromosomal imbalances in advanced urinary bladder cancer detected by comparative genomic hybridization. *Am. J. Pathol.* 153, 1615–1621.
- Bruch, J., Wöhr, G., Hautmann, R., Mattfeldt, T., Bruderlein, S., Moller, P., Sauter, S., Hameister, H., Vogel, W., and Paiss, T. (1998) Chromosomal changes during progression of transitional cell carcinoma of the bladder and delineation of the amplified interval on chromosome arm 8q. *Genes Chromosomes Cancer* 23, 167–174.
- Hovey, R. M., Chu, L., Balazs, M., De Vries, S., Moore, D., Sauter, G., Carroll, P. R., and Waldman, F. M. (1998) Genetic alterations in primary bladder cancers and their metastases. *Cancer Res.* 58, 3555–3560.
- Simon, R., Burger, H., Brinkschmidt, C., Bocker, W., Hertle, L., and Terpe, H. J. (1998) Chromosomal aberrations associated with invasion in papillary superficial bladder cancer. *J. Pathol.* 185, 345–351.
- Koo, S. H., Kwon, K. C., Ihm, C. H., Jeon, Y. M., Park, J. W., and Sul, C. K. (1999) Detection of genetic alterations in bladder tumors by comparative genomic hybridization and cytogenetic analysis. *Cancer Genet. Cytogenet.* 110, 87–93.
- Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 15, 1359–1367.
- Christensen, M., Sunde, L., Bolund, L., and Ørntoft, T. F. (1999) Comparison of three methods of microsatellite detection. *Scand. J. Clin. Lab. Invest.* 59, 167–177.
- Celis, J. E., Østergaard, M., Basse, B., Celis, A., Lauridsen, J. B., Ratz, G. P., Andersen, I., Hein, B., Wolf, H., Ørntoft, T. F., and Rasmussen, H. H. (1996) Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. *Cancer Res.* 56, 4782–4790.
- Celis, J. E., Ratz, G., Basse, B., Lauridsen, J. B., and Celis, A. (1994) In *Cell Biology: A Laboratory Handbook* (Celis, J. E., ed) Vol. 3, pp. 222–230, Academic Press, Orlando, FL.
- Ohlsson, R., Tycko, B., and Sapienza, C. (1998) Monoallelic expression: 'there can only be one'. *Trends Genet.* 14, 435–438.
- Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K., and Burakoff, S. J. (1998) Monoallelic expression of the interleukin-2 locus. *Science* 279, 2118–2121.
- Brannan, C. I., and Bartolomei, M. S. (1999) Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.* 9, 164–170.
- Ohlsson, R., Cui, H., He, L., Pfeifer, S., Malmikumpu, H., Jiang, S., Feinberg, A. P., and Hedborg, F. (1999) Mosaic allelic insulin-like growth factor 2 expression patterns reveal a link between Wilms' tumorigenesis and epigenetic heterogeneity. *Cancer Res.* 59, 3889–3892.
- Cui, H., Hedborg, F., He, L., Nordenskjöld, A., Sandstedt, B., Pfeifer-Ohlsson, S., and Ohlsson, R. (1997) Inactivation of H19, an imprinted and putative tumor repressor gene, is a preneoplastic event during Wilms' tumorigenesis. *Cancer Res.* 57, 4469–4473.
- Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S., and Fink, G. R. (1999) Ploidy regulation of gene expression. *Science* 285, 251–254.
- Tsao, J., Yatabe, Y., Mark, I. D., Haiyan, K., Jones, P. A., and Shibata, D. (2000) Bladder cancer genotype stability during clinical progression. *Genes Chromosomes Cancer* 29, 26–32.
- Zong, Q., Schummer, M., Hood, L., and Morris, D. R. (1999) Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10632–10636.
- Anderson, L., and Seilhamer, J. (1997) Comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 18, 533–537.
- Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292, 929–934.

# Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors

Jonathan R. Pollack<sup>\*,†</sup>, Therese Sørli<sup>§</sup>, Charles M. Perou<sup>¶</sup>, Christian A. Rees<sup>||,††</sup>, Stefanie S. Jeffrey<sup>††</sup>, Per E. Lønning<sup>‡‡</sup>, Robert Tibshirani<sup>§§</sup>, David Botstein<sup>||</sup>, Anne-Lise Børresen-Dale<sup>§</sup>, and Patrick O. Brown<sup>††¶¶</sup>

Departments of <sup>\*</sup>Pathology, <sup>¶</sup>Genetics, <sup>††</sup>Surgery, <sup>§§</sup>Health Research and Policy, and <sup>¶¶</sup>Biochemistry, and <sup>†</sup>Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305; <sup>§</sup>Department of Genetics, Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway; <sup>‡‡</sup>Department of Medicine (Oncology), Haukeland University Hospital, N-5021 Bergen, Norway; and <sup>||</sup>Department of Genetics and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599

Contributed by Patrick O. Brown, August 6, 2002

Genomic DNA copy number alterations are key genetic events in the development and progression of human cancers. Here we report a genome-wide microarray comparative genomic hybridization (array CGH) analysis of DNA copy number variation in a series of primary human breast tumors. We have profiled DNA copy number alteration across 6,691 mapped human genes, in 44 predominantly advanced, primary breast tumors and 10 breast cancer cell lines. While the overall patterns of DNA amplification and deletion corroborate previous cytogenetic studies, the high-resolution (gene-by-gene) mapping of amplicon boundaries and the quantitative analysis of amplicon shape provide significant improvement in the localization of candidate oncogenes. Parallel microarray measurements of mRNA levels reveal the remarkable degree to which variation in gene copy number contributes to variation in gene expression in tumor cells. Specifically, we find that 62% of highly amplified genes show moderately or highly elevated expression, that DNA copy number influences gene expression across a wide range of DNA copy number alterations (deletion, low-, mid- and high-level amplification), that on average, a 2-fold change in DNA copy number is associated with a corresponding 1.5-fold change in mRNA levels, and that overall, at least 12% of all the variation in gene expression among the breast tumors is directly attributable to underlying variation in gene copy number. These findings provide evidence that widespread DNA copy number alteration can lead directly to global deregulation of gene expression, which may contribute to the development or progression of cancer.

Conventional cytogenetic techniques, including comparative genomic hybridization (CGH) (1), have led to the identification of a number of recurrent regions of DNA copy number alteration in breast cancer cell lines and tumors (2–4). While some of these regions contain known or candidate oncogenes [e.g., FGFR1 (8p11), MYC (8q24), CCND1 (11q13), ERBB2 (17q12), and ZNF217 (20q13)] and tumor suppressor genes [RB1 (13q14) and TP53 (17p13)], the relevant gene(s) within other regions (e.g., gain of 1q, 8q22, and 17q22–24, and loss of 8p) remain to be identified. A high-resolution genome-wide map, delineating the boundaries of DNA copy number alterations in tumors, should facilitate the localization and identification of oncogenes and tumor suppressor genes in breast cancer. In this study, we have created such a map, using array-based CGH (5–7) to profile DNA copy number alteration in a series of breast cancer cell lines and primary tumors.

An unresolved question is the extent to which the widespread DNA copy number changes that we and others have identified in breast tumors alter expression of genes within involved regions. Because we had measured mRNA levels in parallel in the same samples (8), using the same DNA microarrays, we had an opportunity to explore on a genomic scale the relationship between DNA copy number changes and gene expression. From

this analysis, we have identified a significant impact of widespread DNA copy number alteration on the transcriptional programs of breast tumors.

## Materials and Methods

**Tumors and Cell Lines.** Primary breast tumors were predominantly large (>3 cm), intermediate-grade, infiltrating ductal carcinomas, with more than 50% being lymph node positive. The fraction of tumor cells within specimens averaged at least 50%. Details of individual tumors have been published (8, 9), and are summarized in Table 1, which is published as supporting information on the PNAS web site, [www.pnas.org](http://www.pnas.org). Breast cancer cell lines were obtained from the American Type Culture Collection. Genomic DNA was isolated either using Qiagen genomic DNA columns, or by phenol/chloroform extraction followed by ethanol precipitation.

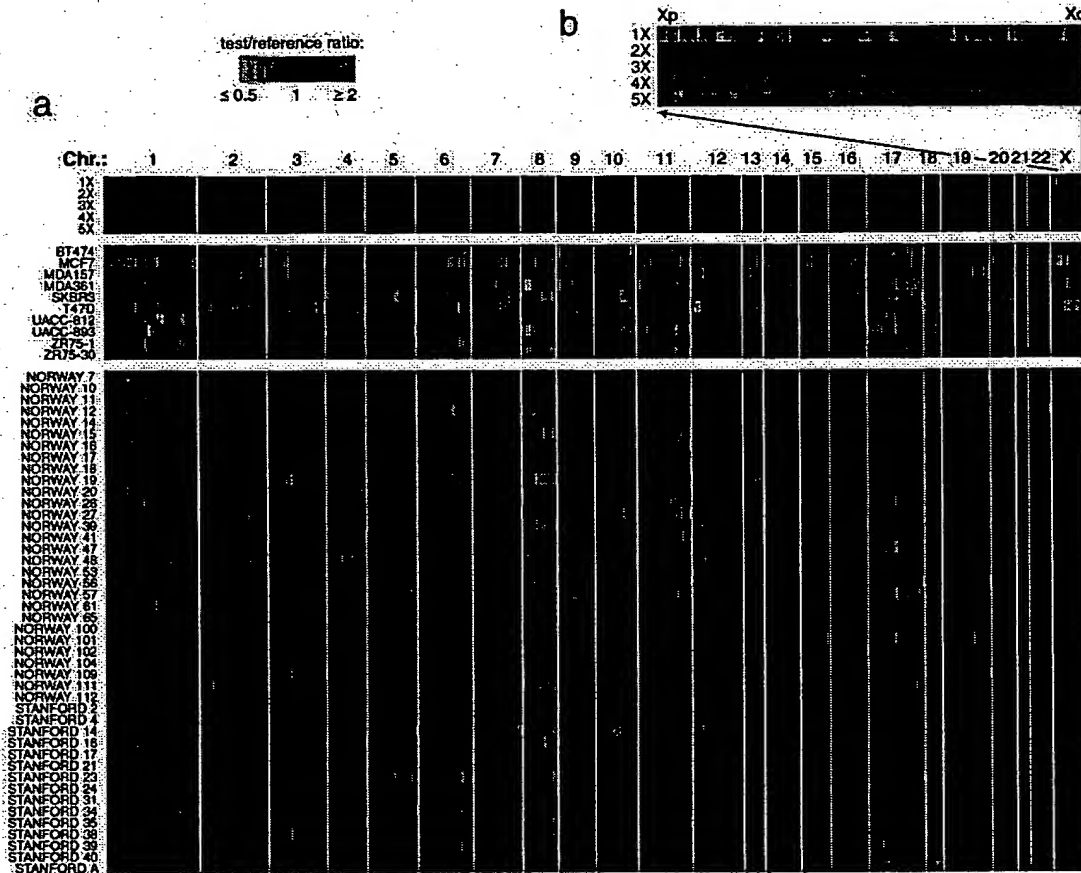
**DNA Labeling and Microarray Hybridizations.** Genomic DNA labeling and hybridizations were performed essentially as described in Pollack *et al.* (7), with slight modifications. Two micrograms of DNA was labeled in a total volume of 50 microliters and the volumes of all reagents were adjusted accordingly. “Test” DNA (from tumors and cell lines) was fluorescently labeled (Cy5) and hybridized to a human cDNA microarray containing 6,691 different mapped human genes (i.e., UniGene clusters). The “reference” (labeled with Cy3) for each hybridization was normal female leukocyte DNA from a single donor. The fabrication of cDNA microarrays and the labeling and hybridization of mRNA samples have been described (8).

**Data Analysis and Map Positions.** Hybridized arrays were scanned on a GenePix scanner (Axon Instruments, Foster City, CA), and fluorescence ratios (test/reference) calculated using SCANALYZE software (available at <http://rana.lbl.gov>). Fluorescence ratios were normalized for each array by setting the average log fluorescence ratio for all array elements equal to 0. Measurements with fluorescence intensities more than 20% above background were considered reliable. DNA copy number profiles that deviated significantly from background ratios measured in normal genomic DNA control hybridizations were interpreted as evidence of real DNA copy number alteration (see *Estimating Significance of Altered Fluorescence Ratios* in the supporting information). When indicated, DNA copy number profiles are displayed as a moving average (symmetric 5-nearest neighbors). Map positions for arrayed human cDNAs were assigned by

Abbreviation: CGH, comparative genomic hybridization.

<sup>\*</sup>To whom reprint requests should be addressed at: Department of Pathology, Stanford University School of Medicine, CCSR Building, Room 3245A, 269 Campus Drive, Stanford, CA 94305-5176. E-mail: [pollack1@stanford.edu](mailto:pollack1@stanford.edu).

<sup>††</sup>Present address: Zyomyx Inc., Hayward, CA 94545.



**Fig. 1.** Genome-wide measurement of DNA copy number alteration by array CGH. (a) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Each row represents a different cell line or tumor, and each column represents one of 6,691 different mapped human genes present on the microarray, ordered by genome map position from 1pter through Xqter. Moving average (symmetric 5-nearest neighbors) fluorescence ratios (test/reference) are depicted using a log<sub>2</sub>-based pseudocolor scale (indicated), such that red luminescence reflects fold-amplification, green luminescence reflects fold-deletion, and black indicates no change (gray indicates poorly measured data). (b) Enlarged view of DNA copy number profiles across the X chromosome, shown for cell lines containing different numbers of X chromosomes.

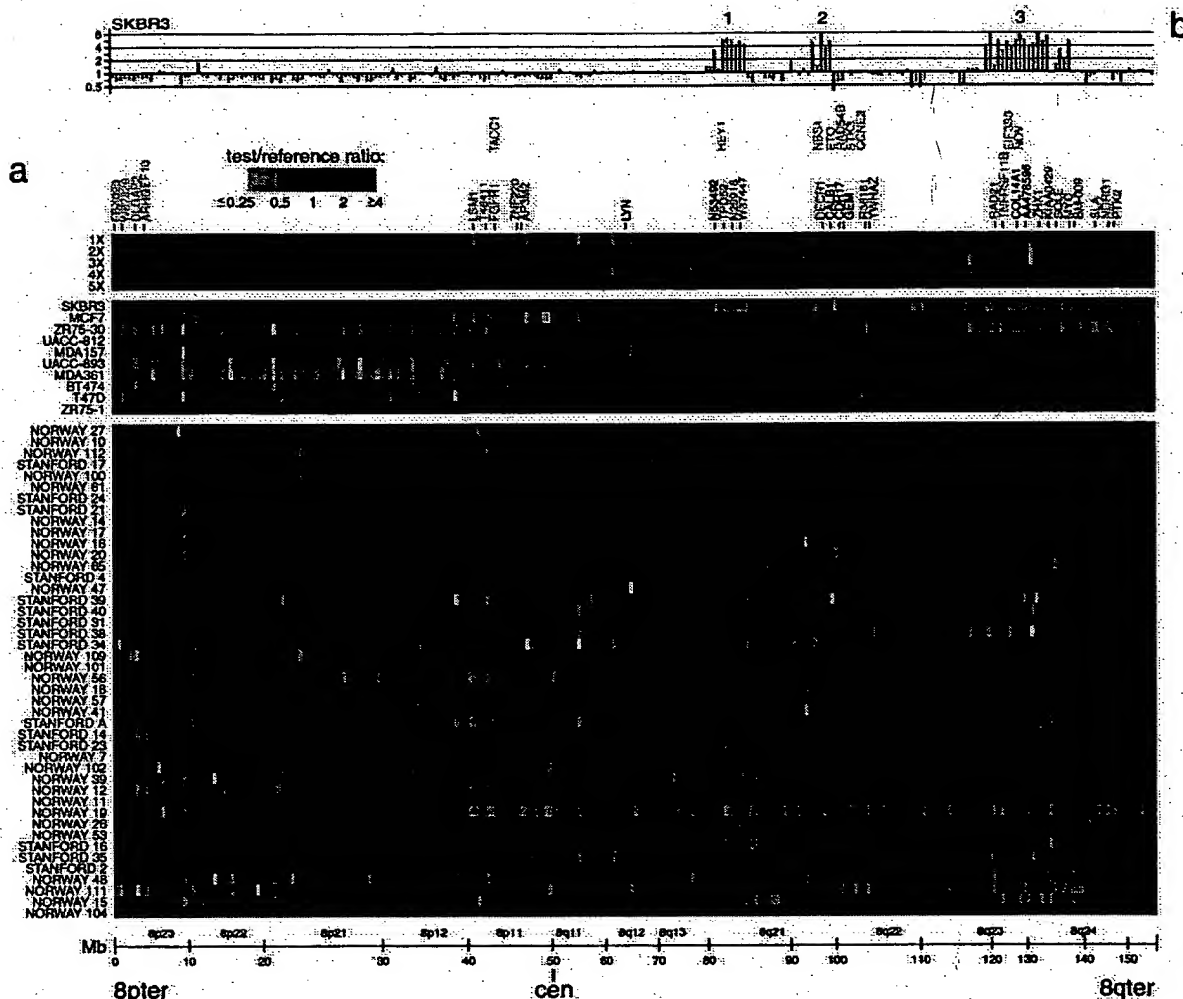
identifying the starting position of the best and longest match of any DNA sequence represented in the corresponding UniGene cluster (10) against the "Golden Path" genome assembly (<http://genome.ucsc.edu/>; Oct 7, 2000 Freeze). For UniGene clusters represented by multiple arrayed elements, mean fluorescence ratios (for all elements representing the same UniGene cluster) are reported. For mRNA measurements, fluorescence ratios are "mean-centered" (i.e., reported relative to the mean ratio across the 44 tumor samples). The data set described here can be accessed in its entirety in the supporting information.

## Results

We performed CGH on 44 predominantly locally advanced, primary breast tumors and 10 breast cancer cell lines, using cDNA microarrays containing 6,691 different mapped human genes (Fig. 1a; also see *Materials and Methods* for details of microarray hybridizations). To take full advantage of the improved spatial resolution of array CGH, we ordered (fluorescence ratios for) the 6,691 cDNAs according to the "Golden Path" (<http://genome.ucsc.edu/>) genome assembly of the draft human genome sequences (11). In so doing, arrayed cDNAs not only themselves represent genes of potential interest (e.g., candidate oncogenes within amplicons), but also provide precise genetic landmarks for chromosomal regions of amplification and

deletion. Parallel analysis of DNA from cell lines containing different numbers of X chromosomes (Fig. 1b), as we did before (7), demonstrated the sensitivity of our method to detect single-copy loss (45, XO), and 1.5- (47, XXX), 2- (48, XXXX), or 2.5-fold (49, XXXXX) gains (also see Fig. 5, which is published as supporting information on the PNAS web site). Fluorescence ratios were linearly proportional to copy number ratios, which were slightly underestimated, in agreement with previous observations (7). Numerous DNA copy number alterations were evident in both the breast cancer cell lines and primary tumors (Fig. 1a), detected in the tumors despite the presence of euploid non-tumor cell types; the magnitudes of the observed changes were generally lower in the tumor samples. DNA copy-number alterations were found in every cancer cell line and tumor, and on every human chromosome in at least one sample. Recurrent regions of DNA copy number gain and loss were readily identifiable. For example, gains within 1q, 8q, 17q, and 20q were observed in a high proportion of breast cancer cell lines/tumors (90%/69%, 100%/47%, 100%/60%, and 90%/44%, respectively), as were losses within 1p, 3p, 8p, and 13q (80%/24%, 80%/22%, 80%/22%, and 70%/18%, respectively), consistent with published cytogenetic studies (refs. 2-4; a complete listing of gains/losses is provided in Tables 2 and 3, which are published as supporting information on the PNAS web site). The total





**Fig. 2.** DNA copy number alteration across chromosome 8 by array CGH. (a) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering to highlight recurrent copy number changes. The 241 genes present on the microarrays and mapping to chromosome 8 are ordered by position along the chromosome. Fluorescence ratios (test/reference) are depicted by a  $\log_2$  pseudocolor scale (indicated). Selected genes are indicated with color-coded text (red, increased; green, decreased; black, no change; gray, not well measured) to reflect correspondingly altered mRNA levels (observed in the majority of the subset of samples displaying the DNA copy number change). The map positions for genes of interest that are not represented on the array are indicated in the row above those genes represented on the array. (b) Graphical display of DNA copy number profile for breast cancer cell line SKBR3. Fluorescence ratios (tumor/normal) are plotted on a  $\log_2$  scale for chromosome 8 genes, ordered along the chromosome.

number of genomic alterations (gains and losses) was found to be significantly higher in breast tumors that were high grade ( $P = 0.008$ ), consistent with published CGH data (3), estrogen receptor negative ( $P = 0.04$ ), and harboring TP53 mutations ( $P = 0.0006$ ) (see Table 4, which is published as supporting information on the PNAS web site).

The improved spatial resolution of our array CGH analysis is illustrated for chromosome 8, which displayed extensive DNA copy number alteration in our series. A detailed view of the variation in the copy number of 241 genes mapping to chromosome 8 revealed multiple regions of recurrent amplification; each of these potentially harbors a different known or previously uncharacterized oncogene (Fig. 2a). The complexity of amplicon structure is most easily appreciated in the breast cancer cell line SKBR3. Although a conventional CGH analysis of 8q in SKBR3 identified only two distinct regions of amplification (12), we observed three distinct regions of high-level amplification (labeled 1–3 in Fig. 2b). For each of these regions we can define the

boundaries of the interval recurrently amplified in the tumors we examined; in each case, known or plausible candidate oncogenes can be identified (a description of these regions, as well as the recurrently amplified regions on chromosomes 17 and 20, can be found in Figs. 6 and 7, which are published as supporting information on the PNAS web site).

For a subset of breast cancer cell lines and tumors (4 and 37, respectively), and a subset of arrayed genes (6,095), mRNA levels were quantitatively measured in parallel by using cDNA microarrays (8). The parallel assessment of mRNA levels is useful in the interpretation of DNA copy number changes. For example, the highly amplified genes that are also highly expressed are the strongest candidate oncogenes within an amplicon. Perhaps more significantly, our parallel analysis of DNA copy number changes and mRNA levels provides us the opportunity to assess the global impact of widespread DNA copy number alteration on gene expression in tumor cells.

A strong influence of DNA copy number on gene expression is evident in an examination of the pseudocolor representations

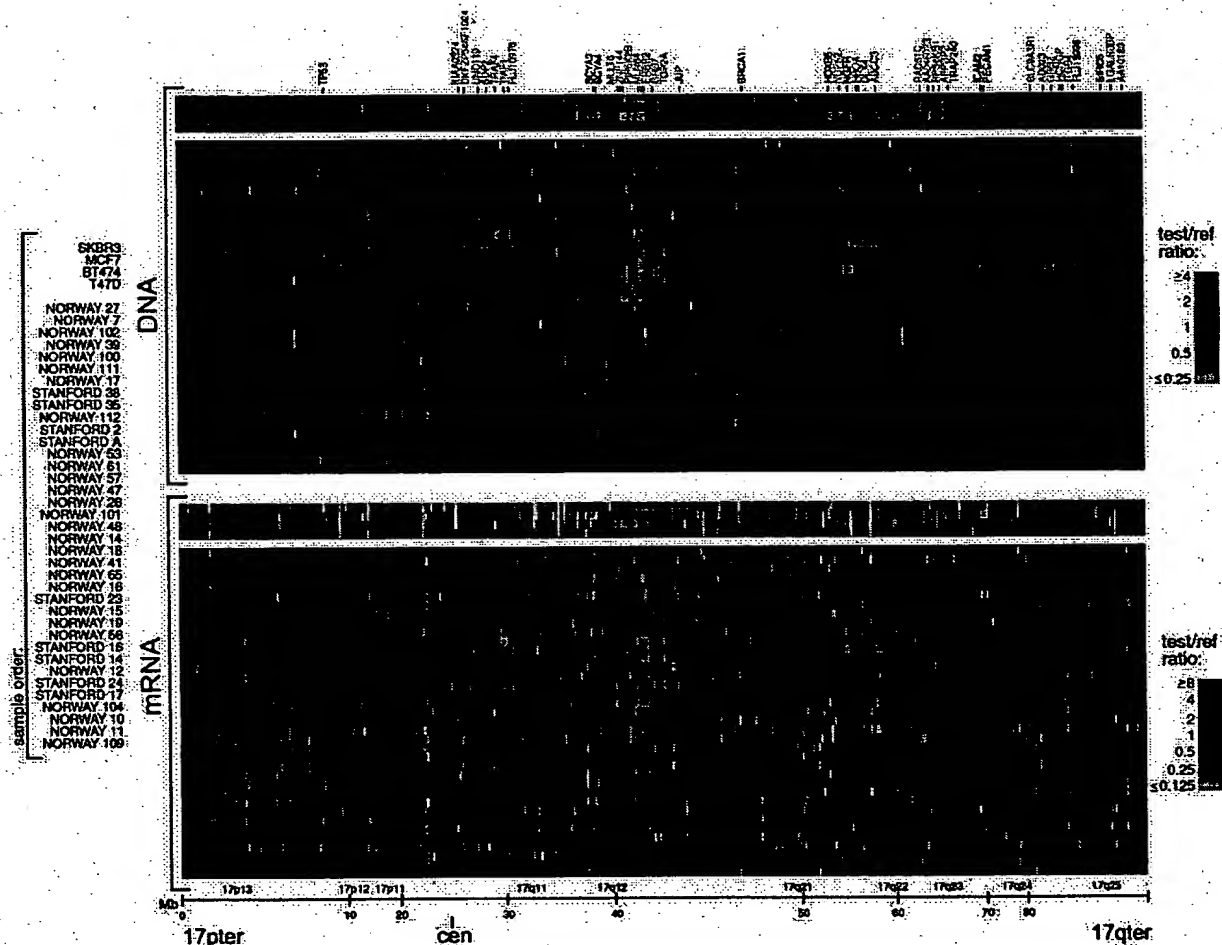


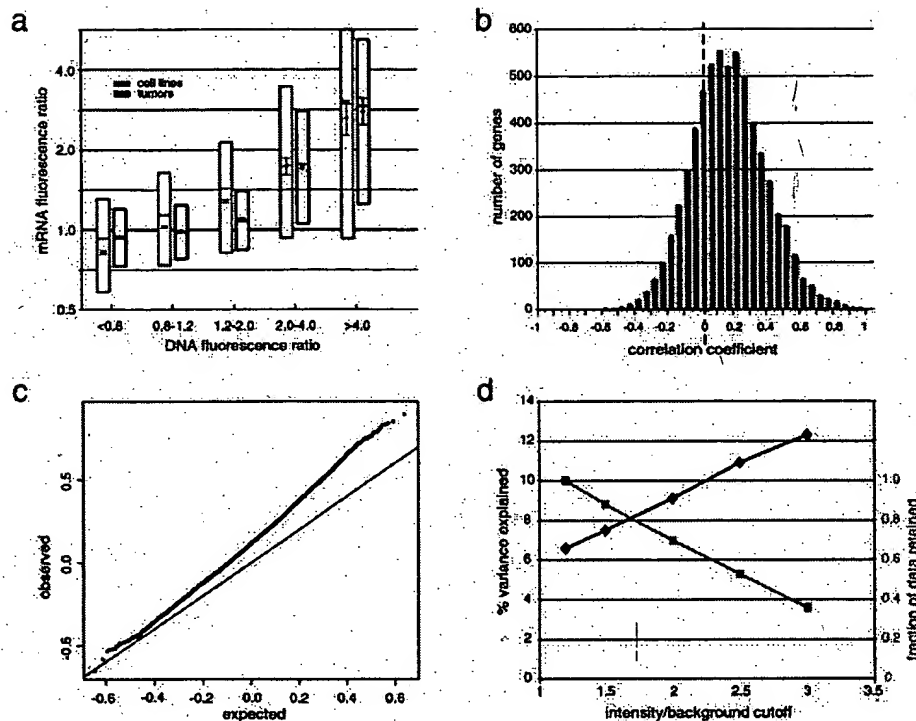
Fig. 3. Concordance between DNA copy number and gene expression across chromosome 17. DNA copy number alteration (Upper) and mRNA levels (Lower) are illustrated for breast cancer cell lines and tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering (Upper), and the identical sample order is maintained (Lower). The 354 genes present on the microarrays and mapping to chromosome 17, and for which both DNA copy number and mRNA levels were determined, are ordered by position along the chromosome; selected genes are indicated in color-coded text (see Fig. 2 legend). Fluorescence ratios (test/reference) are depicted by separate  $\log_2$  pseudocolor scales (indicated).

of DNA copy number and mRNA levels for genes on chromosome 17 (Fig. 3). The overall patterns of gene amplification and elevated gene expression are quite concordant; i.e., a significant fraction of highly amplified genes appear to be correspondingly highly expressed. The concordance between high-level amplification and increased gene expression is not restricted to chromosome 17. Genome-wide, of 117 high-level DNA amplifications (fluorescence ratios  $>4$ , and representing 91 different genes), 62% (representing 54 different genes; see Table 5, which is published as supporting information on the PNAS web site) are found associated with at least moderately elevated mRNA levels (mean-centered fluorescence ratios  $>2$ ), and 42% (representing 36 different genes) are found associated with comparably highly elevated mRNA levels (mean-centered fluorescence ratios  $>4$ ).

To determine the extent to which DNA deletion and lower-level amplification (in addition to high-level amplification) are also associated with corresponding alterations in mRNA levels, we performed three separate analyses on the complete data set (4 cell lines and 37 tumors, across 6,095 genes). First, we determined the average mRNA levels for each of five classes of genes, representing DNA deletion, no change, and low-, medium-, and high-level amplification (Fig. 4a). For both the

breast cancer cell lines and tumors, average mRNA levels tracked with DNA copy number across all five classes, in a statistically significant fashion ( $P$  values for pair-wise Student's  $t$  tests comparing adjacent classes: cell lines,  $4 \times 10^{-49}$ ,  $1 \times 10^{-49}$ ,  $5 \times 10^{-5}$ ,  $1 \times 10^{-2}$ ; tumors,  $1 \times 10^{-43}$ ,  $1 \times 10^{-214}$ ,  $5 \times 10^{-41}$ ,  $1 \times 10^{-4}$ ). A linear regression of the average  $\log(\text{DNA copy number})$ , for each class, against average  $\log(\text{mRNA level})$  demonstrated that on average, a 2-fold change in DNA copy number was accompanied by 1.4- and 1.5-fold changes in mRNA level for the breast cancer cell lines and tumors, respectively (Fig. 4a, regression line not shown). Second, we characterized the distribution of the 6,095 correlations between DNA copy number and mRNA level, each across the 37 tumor samples (Fig. 4b). The distribution of correlations forms a normal-shaped curve, but with the peak markedly shifted in the positive direction from zero. This shift is statistically significant, as evidenced in a plot of observed vs. expected correlations (Fig. 4c), and reflects a pervasive global influence of DNA copy number alterations on gene expression. Notably, the highest correlations between DNA copy number and mRNA level (the right tail of the distribution in Fig. 4b) comprise both amplified and deleted genes (data not shown). Third, we used a linear regression model to estimate the fraction of all variation measured in mRNA levels among the 37





**Fig. 4.** Genome-wide influence of DNA copy number alterations on mRNA levels. (a) For breast cancer cell lines (gray) and tumor samples (black), both mean-centered mRNA fluorescence ratio (log<sub>2</sub> scale) quartiles (box plots indicate 25th, 50th, and 75th percentile) and averages (diamonds; Y-value error bars indicate standard errors of the mean) are plotted for each of five classes of genes, representing DNA deletion (tumor/normal ratio <0.8), no change (0.8–1.2), low- (1.2–2), medium- (2–4), and high-level (>4) amplification. *P* values for pair-wise Student's *t* tests, comparing averages between adjacent classes (moving left to right), are  $4 \times 10^{-49}$ ,  $1 \times 10^{-49}$ ,  $5 \times 10^{-5}$ ,  $1 \times 10^{-2}$  (cell lines), and  $1 \times 10^{-43}$ ,  $1 \times 10^{-214}$ ,  $5 \times 10^{-41}$ ,  $1 \times 10^{-4}$  (tumors). (b) Distribution of correlations between DNA copy number and mRNA levels, for 6,095 different human genes across 37 breast tumor samples. (c) Plot of observed versus expected correlation coefficients. The expected values were obtained by randomization of the sample labels in the DNA copy number data set. The line of unity is indicated. (d) Percent variance in gene expression (among tumors) directly explained by variation in gene copy number. Percent variance explained (black line) and fraction of data retained (gray line) are plotted for different fluorescence intensity/background (a rough surrogate for signal/noise) cutoff values. Fraction of data retained is relative to the 1.2 intensity/background cutoff. Details of the linear regression model used to estimate the fraction of variation in gene expression attributable to underlying DNA copy number alteration can be found in the supporting information (see *Estimating the Fraction of Variation in Gene Expression Attributable to Underlying DNA Copy Number Alteration*).

tumors that could be attributed to underlying variation in DNA copy number. From this analysis, we estimate that, overall, about 7% of all of the observed variation in mRNA levels can be explained directly by variation in copy number of the altered genes (Fig. 4d). We can reduce the effects of experimental measurement error on this estimate by using only that fraction of the data most reliably measured (fluorescence intensity/background >3); using that data, our estimate of the percent variation in mRNA levels directly attributed to variation in gene copy number increases to 12% (Fig. 4d). This still undoubtedly represents a significant underestimate, as the observed variation in global gene expression is affected not only by true variation in the expression programs of the tumor cells themselves, but also by the variable presence of non-tumor cell types within clinical samples.

## Discussion

This genome-wide, array CGH analysis of DNA copy number alteration in a series of human breast tumors demonstrates the usefulness of defining amplicon boundaries at high resolution (gene-by-gene), and quantitatively measuring amplicon shape, to assist in locating and identifying candidate oncogenes. By analyzing mRNA levels in parallel, we have also discovered that changes in DNA copy number have a large, pervasive, direct effect on global gene expression patterns in both breast cancer

cell lines and tumors. Although the DNA microarrays used in our analysis may display a bias toward characterized and/or highly expressed genes, because we are examining such a large fraction of the genome (approximately 20% of all human genes), and because, as detailed above, we are likely underestimating the contribution of DNA copy number changes to altered gene expression, we believe our findings are likely to be generalizable (but would nevertheless still be remarkable if only applicable to this set of ~6,100 genes).

In budding yeast, aneuploidy has been shown to result in chromosome-wide gene expression biases (13). Two recent studies have begun to examine the global relationship between DNA copy number and gene expression in cancer cells. In agreement with our findings, Phillips *et al.* (14) have shown that with the acquisition of tumorigenicity in an immortalized prostate epithelial cell line, new chromosomal gains and losses resulted in a statistically significant respective increase and decrease in the average expression level of involved genes. In contrast, Platzer *et al.* (15) recently reported that in metastatic colon tumors only ~4% of genes within amplified regions were found more highly (>2-fold) expressed, when compared with normal colonic epithelium. This report differs substantially from our finding that 62% of highly amplified genes in breast cancer exhibit at least 2-fold increased expression. These contrasting findings may reflect methodological differences between the

studies. For example, the study of Platzter *et al.* (15) may have systematically under-measured gene expression changes. In this regard it is remarkable that only 14 transcripts of many thousand residing within unamplified chromosomal regions were found to exhibit at least 4-fold altered expression in metastatic colon cancer. Additionally, their reliance on lower-resolution chromosomal CGH may have resulted in poorly delimiting the boundaries of high-complexity amplicons, effectively overcalling regions with amplification. Alternatively, the contrasting findings for amplified genes may represent real biological differences between breast and metastatic colon tumors; resolution of this issue will require further studies.

Our finding that widespread DNA copy number alteration has a large, pervasive and direct effect on global gene expression patterns in breast cancer has several important implications. First, this finding supports a high degree of copy number-dependent gene expression in tumors. Second, it suggests that most genes are not subject to specific autoregulation or dosage compensation. Third, this finding cautions that elevated expression of an amplified gene cannot alone be considered strong independent evidence of a candidate oncogene's role in tumorigenesis. In our study, fully 62% of highly amplified genes demonstrated moderately or highly elevated expression. This highlights the importance of high-resolution mapping of amplicon boundaries and shape [to identify the "driving" gene(s) within amplicons (16)], on a large number of samples, in addition to functional studies. Fourth, this finding suggests that analyzing

the genomic distribution of expressed genes, even within existing microarray gene expression data sets, may permit the inference of DNA copy number aberration, particularly aneuploidy (where gene expression can be averaged across large chromosomal regions; see Fig. 3 and supporting information). Fifth, this finding implies that a substantial portion of the phenotypic uniqueness (and by extension, the heterogeneity in clinical behavior) among patients' tumors may be traceable to underlying variation in DNA copy number. Sixth, this finding supports a possible role for widespread DNA copy number alteration in tumorigenesis (17, 18), beyond the amplification of specific oncogenes and deletion of specific tumor suppressor genes. Widespread DNA copy number alteration, and the concomitant widespread imbalance in gene expression, might disrupt critical stoichiometric relationships in cell metabolism and physiology (e.g., proteasome, mitotic spindle), possibly promoting further chromosomal instability and directly contributing to tumor development or progression. Finally, our findings suggest the possibility of cancer therapies that exploit specific or global imbalances in gene expression in cancer.

We thank the many members of the P.O.B. and D.B. labs for helpful discussions. J.R.P. was a Howard Hughes Medical Institute Physician Postdoctoral Fellow during a portion of this work. P.O.B. is a Howard Hughes Medical Institute Associate Investigator. This work was supported by grants from the National Institutes of Health, the Howard Hughes Medical Institute, the Norwegian Cancer Society, and the Norwegian Research Council.

1. Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rütovitz, D., Gray, J. W., Waldman, F. & Pinkel, D. (1992) *Science* 258, 818–821.
2. Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W. & Waldman, F. M. (1994) *Proc. Natl. Acad. Sci. USA* 91, 2156–2160.
3. Tirkkonen, M., Tanner, M., Karhu, R., Kallioniemi, A., Isola, J. & Kallioniemi, O. P. (1998) *Genes Chromosomes Cancer* 21, 177–184.
4. Forozan, F., Mahlamaki, E. H., Monni, O., Chen, Y., Veldman, R., Jiang, Y., Gooden, G. C., Ethier, S. P., Kallioniemi, A. & Kallioniemi, O. P. (2000) *Cancer Res.* 60, 4519–4525.
5. Solinas-Toldo, S., Lampel, S., Stülgemeyer, S., Nickolenko, J., Benner, A., Döhner, H., Cremer, T. & Lichter, P. (1997) *Genes Chromosomes Cancer* 20, 399–407.
6. Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., *et al.* (1998) *Nat. Genet.* 20, 207–211.
7. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) *Nat. Genet.* 23, 41–46.
8. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* (2000) *Nature (London)* 406, 747–752.
9. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98, 10869–10874.
10. Schuler, G. D. (1997) *J. Mol. Med.* 75, 694–698.
11. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001) *Nature (London)* 409, 860–921.
12. Fejzo, M. S., Godfrey, T., Chen, C., Waldman, F. & Gray, J. W. (1998) *Genes Chromosomes Cancer* 22, 105–113.
13. Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., *et al.* (2000) *Nat. Genet.* 25, 333–337.
14. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., *et al.* (2001) *Cancer Res.* 61, 8143–8149.
15. Platzter, P., Upender, M. B., Wilson, K., Willis, J., Lutterbaugh, J., Nosrati, A., Willson, J. K., Mack, D., Ried, T. & Markowitz, S. (2002) *Cancer Res.* 62, 1134–1138.
16. Albertson, D. G., Ylstra, B., Segraves, R., Collins, C., Dairkee, S. H., Kowbel, D., Kuo, W. L., Gray, J. W. & Pinkel, D. (2000) *Nat. Genet.* 25, 144–146.
17. Li, R., Yerganian, G., Duesberg, P., Kraemer, A., Willer, A., Rausch, C. & Hehlmann, R. (1997) *Proc. Natl. Acad. Sci. USA* 94, 14506–14511.
18. Rasnick, D. & Duesberg, P. H. (1999) *Biochem. J.* 340, 621–630.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**